



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:04:29 ; Search time 93.7059 Seconds  
(without alignments)  
1002.955 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303  
Sequence: 1 MRPGPAPASQRLRGILL.....GDASTGMSVSRILIELPK 243

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1980s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	AAB36667	Aab36667 Human sec
2	1303	100.0	243	ABG96340	Abg96340 Human ova
3	1303	100.0	243	ABJ05554	Abj05554 Breast ca
4	1303	100.0	243	ABB80978	Abb80978 Human REM
5	1303	100.0	243	ABR58546	AbR58546 Human can
6	1303	100.0	243	ABR48227	AbR48227 Human bla
7	1303	100.0	243	ABG75758	AbG75758 Human REM
8	1303	100.0	243	ABU56607	Abu56607 Lung canc
9	1303	100.0	243	ADB80510	AdB80510 Ovarian c
10	1303	100.0	243	ADN39855	Adn39855 Cancer/an
11	1303	100.0	243	ADN38732	AdN38732 Cancer/an
12	1303	100.0	243	ADL70254	AdL70254 IBFLJ01 P
13	1303	100.0	278	AAB08856	Aab08856 Amino aci
14	1303	100.0	278	AAM25746	Aam25746 Human pro
15	1303	100.0	278	ABG96338	AbG96338 Human ova
16	1303	100.0	278	ABR47627	AbR47627 Breast ca
17	1300	99.8	243	AAV91529	Aay91529 Human sec
18	1300	99.8	278	ADL71601	AdL71601 Novel hum
19	1300	99.8	243	ABP66631	Abp66631 Human pan
20	1298	99.6	243	ABJ37031	Abj37031 Human bre
21	1294	99.3	243	AAV99462	Aay99462 Human PRO
22	1294	99.3	243	AAV99441	Aay99441 Human PRO
23	1294	99.3	243	AAB66190	Aab66190 Protein o
24	1294	99.3	243	AAB66211	Aab66211 Protein o
25	1294	99.3	243	AAU29206	Aau29206 Human PRO

#### ALIGNMENTS

26	1294	99.3	243	4	AAB87586	Aab87586 Human PRO
27	1294	99.3	243	5	ABG95911	Abg95911 Human sec
28	1294	99.3	243	5	ABB84939	Abb84939 Human PRO
29	1294	99.3	243	5	AAE20462	Aae20462 Human tum
30	1294	99.3	243	5	ABR95545	Abb95545 Human ang
31	1294	99.3	243	5	ABP68636	Abp68636 Human pan
32	1294	99.3	243	5	ABG78938	Abg78938 Human bre
33	1294	99.3	243	6	ABU58582	Abu58582 Human PRO
34	1294	99.3	243	6	ABU88130	Abu88130 Novel hum
35	1294	99.3	243	6	ABU84445	Abu84445 Human sec
36	1294	99.3	243	6	ABR66319	AbR66319 Human sec
37	1294	99.3	243	6	ABR65709	AbR65709 Human sec
38	1294	99.3	243	6	ABU96649	Abu96649 Human sec
39	1294	99.3	243	6	ABJ37761	Abj37761 Human tum
40	1294	99.3	243	6	ABU82888	Abu82888 Human PRO
41	1294	99.3	243	6	ABU90009	Abu90009 Novel hum
42	1294	99.3	243	6	ABR68258	AbR68258 Human sec
43	1294	99.3	243	6	ABU96311	Abu96311 Novel hum
44	1294	99.3	243	6	ABU92742	Abu92742 Human sec
45	1294	99.3	243	6	ABO08819	AbO08819 Human sec

RESULT 1  
AAB36667 standard; protein; 243 AA.  
ID AAB36667;  
AC AAB36667;  
XX  
XX  
DT 14-MAR-2001 (first entry)  
XX  
DE Human secretory protein TGC-628 SEQ ID NO:7.  
XX  
XX Human; secretory protein; cancer; immune disease; infectious disease;  
XX lung function disorder; liver function disorder; antiinflammatory;  
XX gastrointestinal disorder; cytostatic; haematopoietic; anticonagulant;  
XX immunomodulatory; hepatotropic; cell proliferation-stimulant;  
XX cell migratory agent; cell differentiation-inducer.  
OS Homo sapiens.  
XX  
XX  
XX PN W0200071581-A1.  
XX  
XX PD 30-NOV-2000.  
XX  
XX PF 19-MAY-2000; 2000WO-JP003221.  
XX  
XX PR 20-MAY-1999; 99JP-00140229.  
XX  
XX PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX PI Itoh Y, Mogi S, Tanaka H, Onkubo S, Ogi K;  
XX WPI; 2001-032023/04.  
XX DR N-PDB; AAC90707.  
XX  
XX Novel secretory protein and its salt with e.g. anti-cancer, anti-  
XX inflammatory and hematopoietic, effects, applicable as drugs in remedies  
XX and preventives to treat diseases like cancer and immune diseases.  
XX  
XX Claim 1; Page 89-90; 122pp; Japanese.  
XX  
XX AAC90701 to AAC90715 encode the human secretory proteins given in  
XX AAB36661 to AAB36675. The proteins can have cytostatic, anti-  
XX inflammatory, haematopoietic, anti-coagulant, immunomodulatory and  
XX hepatotropic activities, and can be used as cell migratory agents, cell  
XX proliferation-stimulants and cell differentiation-inducers. The proteins  
XX are useful in the treatment and prevention of diseases such as cancer,  
XX lung function disorder, liver function disorder, gastrointestinal  
XX disorder and immune diseases. AAC90716 to AAC90755 represent PCR primers  
XX which are used in the exemplification of the present invention

```
XX Sequence 243 AA;
SQ
Query Match      100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLRGLLLLLLLQLPAPSSASIPKQKQAKQREVVLDLYNMGCLQGA 60
    |||||
DB 1 MRPGPASPQRRLRGLLLLLLLQLPAPSSASIPKQKQAKQREVVLDLYNMGCLQGA 60

QY 61 GVPGRDGSFGANGIGTPTGIPGRDGFKEGECLEESFEESWTPTNYKQCSWSLNYGIDL 120
    |||||
DB 61 GVPGRDGSFGANGIGTPTGIPGRDGFKEGECLEESFEESWTPTNYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWYTFNGAECGGLPIEAIIVYDQ 180
    |||||
DB 121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWYTFNGAECGGLPIEAIIVYDQ 180

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
    |||||
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240

QY 241 LPK 243
    |||
DB 241 LPK 243

RESULT 2
ABG96340
ID ABG96340 standard; protein; 243 AA.
XX
AC ABG96340;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker M450.
XX
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
    central nervous system disorder; bacterial meningitis; viral meningitis;
    Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
    brain herniation; inflammation; encephalitis; testicular disorder;
    nondercutulous granulomacous orchitis; connective tissue disorder;
    heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
    histological type; carcinogenic; ovarian cancer marker.
XX
OS Homo sapiens.
XX
PN WO200271928-A2.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US007826.
XX
PR 14-MAR-2001; 2001US-0276025P.
    14-MAR-2001; 2001US-0276026P.
    10-AUG-2001; 2001US-0311732P.
    19-SEP-2001; 2001US-0323580P.
    26-SEP-2001; 2001US-0324967P.
    26-SEP-2001; 2001US-0325102P.
    26-SEP-2001; 2001US-0325149P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovakis SG,
    Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vleby PO, Mills GB,
    Baat RC, Lu K, Schmandt RE, Zhao X, Glatc K;
XX
DR MPI: 2002-723277/78.
XX
DR N-PSDB; AB876436.
XX
PT Assessing whether a patient is afflicted with ovarian cancer, useful in
    assessing the stage or progression of the disease, comprises comparing
```

```
PT the expression level of a cancer marker in a sample from a patient and
    from a non cancer patient.
XX
PS Disclosure; Page 252; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
    patient is afflicted with ovarian cancer. The method involves comparing
    the expression level of a marker in a patient sample and the normal level
    of expression of the marker in a control non-ovarian cancer sample, where
    the marker is selected from 363 cancer markers described in the
    specification. The method of the invention is useful in diagnosing or
    characterising cancer, in detecting the presence of cancer as early as
    possible, and the recurrence of ovarian cancer. The method may also be of
    particular use with patients having an enhanced risk of developing
    ovarian cancer (e.g. patients having a familial history of ovarian
    cancer). The cancer markers may be used in the management and treatment
    of e.g. brain and central nervous system disorders (e.g. bacterial and
    viral meningitis, Alzheimer's disease or Parkinson's disease), brain
    disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
    inflammations (e.g. bacterial or viral meningitis or encephalitis),
    testicular disorders (e.g. nondercutulous granulomacous orchitis),
    connective tissue disorders, or heart disorders (e.g. ischaemic heart
    disease or atherosclerosis). The compositions and methods may also be
    used in assessing the histological type of neoplasm associated with
    ovarian cancer, monitoring the progression of ovarian cancer, determining
    whether ovarian cancer has metastasized or is likely to metastasize,
    selecting a composition for inhibiting ovarian cancer, assessing the
    ovarian carcinogenic potential of a compound, or inhibiting ovarian
    cancer or at risk of developing ovarian cancer. The present amino acid
    sequence represents one of the ovarian cancer markers described in the
    invention
XX
SQ Sequence 243 AA;
Query Match      100.0%; Score 1303; DB 5; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLRGLLLLLLLQLPAPSSASIPKQKQAKQREVVLDLYNMGCLQGA 60
    |||||
DB 1 MRPGPASPQRRLRGLLLLLLLQLPAPSSASIPKQKQAKQREVVLDLYNMGCLQGA 60

QY 61 GVPGRDGSFGANGIGTPTGIPGRDGFKEGECLEESFEESWTPTNYKQCSWSLNYGIDL 120
    |||||
DB 61 GVPGRDGSFGANGIGTPTGIPGRDGFKEGECLEESFEESWTPTNYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWYTFNGAECGGLPIEAIIVYDQ 180
    |||||
DB 121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWYTFNGAECGGLPIEAIIVYDQ 180

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
    |||||
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240

QY 241 LPK 243
    |||
DB 241 LPK 243

RESULT 3
ABJ05554
ID ABJ05554 standard; protein; 243 AA.
XX
AC ABJ05554;
XX
DT 14-NOV-2002 (first entry)
XX
DE Breast cancer-associated protein 19.
XX
KW Breast cancer; breast cancer-associated gene sequence; drug development;
    pharmacogenetics; biosensor development.
XX
OS Unidentified.
```



XX MO200259377-A2.  
 FN  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 24-JAN-2002; 2002MO-US002242.  
 XX  
 PR 24-JAN-2001; 2001US-0263965P.  
 PR 02-FEB-2001; 2001US-0265928P.  
 PR 09-APR-2001; 2001US-00829472.  
 PR 04-MAY-2001; 2001US-0282698P.  
 PR 29-MAY-2001; 2001US-0288590P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Mack DH, Gish KC, Afar D;  
 XX  
 DR WPI; 2002-583738/62.  
 DR N-PSDB; ABT07711.  
 XX  
 PT Detecting a breast cancer-associated transcript in a patient's cell,  
 PT sample with a polynucleotide that selectively hybridizes with breast  
 PT cancer nucleic acids.  
 XX  
 PS Disclosure; Page 364; 414pp; English.  
 XX  
 CC The invention comprises a method of detecting a breast cancer-associated  
 CC transcript in a cell from a patient. The method of the invention involves  
 CC contacting a biological sample from the patient with a nucleotide that  
 CC hybridizes to one of the 69 breast cancer-associated gene sequences shown  
 CC in the specification. The method of the invention is useful in the  
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are  
 CC up or down-regulated in breast cancer cells. Genes identified by the  
 CC method of the invention can be used in diagnostic purposes and also as  
 CC targets for screening for therapeutic compounds that modulate breast  
 CC cancer (e.g. hormones or antibodies). Identification of genes that are  
 CC over or under expressed in breast cancer can additionally provide high-  
 CC resolution, high-sensitivity datasets which can be used in the areas of  
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein  
 CC structure and biosensor development. Amino acid sequences ABT05536 -  
 CC ABT05604 represent the proteins encoded by the 69 breast cancer-  
 CC associated genes of the invention  
 XX  
 XX Sequence 243 AA;  
 SQ  
 Query Match 100.0%; Score 1303; DB 5; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 9e-121;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRPOGPAASPORLRLGILLILLQLPAPSSASSEIPKQKQALRQREVVLDLYNGMCLQGPA 60  
 DB 1 MRPOGPAASPORLRLGILLILLQLPAPSSASSEIPKQKQALRQREVVLDLYNGMCLQGPA 60  
 QY 61 GVGGRDSSPGANGIPGPIGPGDGFGEKGECECLRESESSWPNKYQCSWSSINYGIDL 120  
 DB 61 GVGGRDSSPGANGIPGPIGPGDGFGEKGECECLRESESSWPNKYQCSWSSINYGIDL 120  
 QY 121 GKIAECTFTMRNSALRLVFSGLRLKCRNACQQRVYFFNNAECGSPPIFAIYILDQ 180  
 DB 121 GKIAECTFTMRNSALRLVFSGLRLKCRNACQQRVYFFNNAECGSPPIFAIYILDQ 180  
 QY 181 GSPBNASTINIHTSSVEGLCEGIGAGLVDAIIVGTCSDYPKQDASTGNSVSRITIEE 240  
 DB 181 GSPBNASTINIHTSSVEGLCEGIGAGLVDAIIVGTCSDYPKQDASTGNSVSRITIEE 240  
 QY 241 LPK 243  
 DB 241 LPK 243

ABB80978  
 ID ABB80978 standard; protein; 243 AA.  
 XX  
 AC ABB80978;  
 XX  
 DT 21-OCT-2002 (first entry)  
 XX  
 DE Human REMODELIN polypeptide.  
 XX  
 KW REMODELIN; vulnery; vasotrophic; cyostatic; osteopathic; collagen;  
 KW gene therapy; bone; human.  
 XX  
 OS Homo sapiens.  
 PN  
 PN WO200242487-A2.  
 PD 30-MAY-2002.  
 PF 19-OCT-2001; 2001MO-US050940.  
 PF 19-OCT-2000; 2000US-00692081.  
 PR (MAIN-) MAINE MEDICAL CENT RES INST.  
 PA  
 PI Linder V, Friesel R;  
 XX  
 XX WPI; 2002-590472/63.  
 DR N-PSDB; ABN6481.  
 XX  
 PT New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
 PT negative remodeling, fibrosis, collagen disease, and bone disease in a  
 PT mammal.  
 XX  
 PS Claim 5; Fig 4B; 187pp; English.  
 XX  
 CC The invention relates to mammalian REMODELIN polypeptides and encoding  
 CC polynucleotides. REMODELIN is used to diagnose arterial restenosis,  
 CC negative remodeling or fibrosis, bone disease such as osteogenesis  
 CC imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis  
 CC bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are  
 CC used to affect cellular gene expression in a mammal, where the cellular  
 CC gene is from transforming growth factor (TGF)-beta1, collagen I(alphal,  
 CC osteopontin, biglycan, alkaline phosphatase or bone morphogenetic protein  
 CC 4. The expression of osteopontin is dependent on Cbfa1. Compositions  
 CC comprising antisense REMODELIN sequences are useful for treating diseases  
 CC mediated by abnormal expression of a REMODELIN molecule in a human such  
 CC as impaired wound healing, fibrosis of an organ, ectopic ossification, or  
 CC hypertrophic scar formation. REMODELIN is useful in the development of  
 CC bone during mammalian embryogenesis, proliferation and/or migration.  
 CC cellular signaling, adventitial fibrosis, negative remodeling and arterial  
 CC restenosis, smooth muscle cell proliferation, and arterial remodeling and  
 CC useful in gene therapy. REMODELIN is useful for elucidating the function  
 CC of REMODELIN molecules in a cell, to identify a compound that affects  
 CC REMODELIN expression and/or TGF-beta signaling, as a potential  
 CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
 CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
 CC formation, to promote bone fracture healing, and to increase or decrease  
 CC expression of REMODELIN in mammal. It is also useful for producing  
 CC recombinant cell and transgenic non-human mammals which are useful tools  
 CC for the study of REMODELIN action, for identifying novel diagnostics and  
 CC therapeutics for treatment, and for elucidating the cellular roles of  
 CC REMODELIN. The present sequence represents a human REMODELIN polypeptide  
 XX  
 XX Sequence 243 AA;  
 SQ  
 Query Match 100.0%; Score 1303; DB 5; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 9e-121;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRPOGPAASPORLRLGILLILLQLPAPSSASSEIPKQKQALRQREVVLDLYNGMCLQGPA 60  
 DB 1 MRPOGPAASPORLRLGILLILLQLPAPSSASSEIPKQKQALRQREVVLDLYNGMCLQGPA 60

QY 61 GVPGRDGPANGIPGTGPIGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120  
 |||||  
 Db 61 GVPGRDGPANGIPGTGPIGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120  
 QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIYYLDQ 180  
 |||||  
 Db 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIYYLDQ 180  
 QY 181 GSPENNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDPYKGDASTGMSVSRITIEE 240  
 |||||  
 Db 181 GSPENNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDPYKGDASTGMSVSRITIEE 240  
 QY 241 LPK 243  
 ||||  
 Db 241 LPK 243

RESULT 5  
 ABR58546  
 ID ABR58546 standard; protein; 243 AA.  
 XX  
 AC ABR58546;  
 XX  
 DT 09-JUL-2003 (first entry)  
 XX  
 DE Human cancer related protein SEQ ID NO:203.  
 XX  
 KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;  
 KW heart disease; atherosclerosis; endometritis.  
 OS Homo sapiens.  
 OS  
 PN MO2003025138-A2.  
 XX  
 PD 27-MAR-2003.  
 XX  
 PF 17-SEP-2002; 2002MO-US029560.  
 XX  
 PR 17-SEP-2001; 2001US-0323469P.  
 PR 20-SEP-2001; 2001US-0323887P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 08-FEB-2002; 2002US-0355145P.  
 PR 08-FEB-2002; 2002US-0355257P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;  
 PI Zlotnick A;  
 XX  
 DR N-PSDB; ACC72666.  
 DR WPI: 2003-354600/33.  
 PS Claim 12; Page 736; 767pp; English.  
 CC The present invention describes an isolated nucleic acid molecule, which  
 CC comprises the sequence of any of the genes that are up-regulated or down-  
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in  
 CC acute lymphocytic leukemia). ACC72666 represents cancer  
 CC related gene nucleotide sequences which encode the proteins given in  
 CC ABR58521 to ABR58709. Also described: (1) determining the presence or  
 CC absence of a pathological cell in a patient; (2) an expression vector  
 CC comprising a nucleic acid molecule described above; (3) a host cell  
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by  
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide  
 CC of (4); (6) specifically targeting a compound to a pathological cell in a  
 CC patient by administering to the patient the antibody above; and (7) a  
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or  
 CC therapeutic targets. In particular, the nucleic acid is useful for

CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,  
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,  
 CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,  
 CC atherosclerosis and endometritis. The nucleic acid is also useful in  
 CC drug screening, particularly for identifying agents for treating these  
 CC pathologies  
 XX  
 SO Sequence 243 AA;  
 Query Match 100.0%; Score 1303; DB 6; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 9e-121;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPORLRGLLTLTLQLPAPSSASIPKSKOKAOLROREVDLYNMGCLQCPA 60  
 |||||  
 Db 1 MRPGPASPORLRGLLTLTLQLPAPSSASIPKSKOKAOLROREVDLYNMGCLQCPA 60  
 QY 61 GVPGRDGPANGIPGTGPIGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120  
 |||||  
 Db 61 GVPGRDGPANGIPGTGPIGRDGFKEGKECLRESFEESWTPNYKQCSWSLNYGIDL 120  
 QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIYYLDQ 180  
 |||||  
 Db 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIYYLDQ 180  
 QY 181 GSPENNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDPYKGDASTGMSVSRITIEE 240  
 |||||  
 Db 181 GSPENNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDPYKGDASTGMSVSRITIEE 240  
 QY 241 LPK 243  
 ||||  
 Db 241 LPK 243

RESULT 6  
 ABR48227  
 ID ABR48227 standard; protein; 243 AA.  
 XX  
 AC ABR48227;  
 XX  
 DT 12-JUN-2003 (first entry)  
 XX  
 DE Human bladder cancer associated protein sequence SEQ ID NO:175.  
 XX  
 KW Human; bladder cancer; cytostatic; gene therapy; vaccine.  
 OS Homo sapiens.  
 OS  
 PN MO2003003906-A2.  
 XX  
 PD 16-JAN-2003.  
 XX  
 PF 03-JUL-2002; 2002MO-US021338.  
 XX  
 PR 03-JUL-2001; 2001US-0302814P.  
 PR 03-AUG-2001; 2001US-0310099P.  
 PR 08-NOV-2001; 2001US-0343705P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Mack DH, Aziz N;  
 XX  
 DR WPI: 2003-201532/19.  
 DR N-PSDB; ACC51043.  
 XX  
 PT Detecting a bladder cancer-associated transcript in a cell from a  
 PT patient, comprises contacting a biological sample from the patient with a  
 PT bladder cancer-associated polynucleotide or antibody.  
 XX  
 PS Claim 10; Page 289-290; 307pp; English.  
 XX

CC The present invention describes a method for detecting a bladder cancer-  
CC associated transcript in a cell from a patient. The method comprises  
CC contacting a biological sample from the patient with a polynucleotide  
CC that selectively hybridizes to a sequence that is 80 % identical to a  
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059  
CC encode the human bladder cancer-associated proteins given in ABR48146 to  
CC ABR48242). Bladder cancer-associated sequences from the present invention  
CC have cytostatic activities, and can be used in antisense gene therapy and  
CC in vaccine production. The method can be used for detecting a bladder  
CC cancer-associated transcript in a cell from a patient. The method is  
CC useful in diagnosing or treating bladder cancer and in screening for  
CC compounds that modulate bladder cancer, such as hormones or antibodies.  
CC The nucleic acid molecules from the present invention may be used in  
CC various screening and diagnostic methods, and for gene therapy, vaccine  
CC and/or antisense/inhibition applications

XX Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 9e-121; Mismatches 0; Gaps 0;

Matches 243; Conservative 0; Indels 0; Gaps 0;

QY 1 MRPGPAPSPORLRLGLLLLLQLPAPSSASSEIPKQKQAKLRQREVVDLYNGMCIQGPA 60

1 MRPGPAPSPORLRLGLLLLLQLPAPSSASSEIPKQKQAKLRQREVVDLYNGMCIQGPA 60

QY 61 GVPGRDGSPGANGIPETPGIPGRDGFKEGECLERESFESWTPTYKQCSWSLNYGIDL 120

61 GVPGRDGSPGANGIPETPGIPGRDGFKEGECLERESFESWTPTYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRYTFENGACSGPLPIEAIITYLQ 180

121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRYTFENGACSGPLPIEAIITYLQ 180

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240

181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240

QY 241 LPK 243

241 LPK 243

DB 241 LPK 243

RESULT 7

ABG75758

ID ABG75758 standard; protein; 243 AA.

XX ABG75758;

XX 28-APR-2003 (first entry)

DE Human REMODELIN protein.

XX Human; adventitia-inducible; REMODELIN; REMODEL;

KM adventitia induced bone expressed molecule; AIBE; TGF-beta;

KM transforming growth factor beta; adventitia; vascular remodelling;

KM restenosis; vascular injury; antisense therapy; TGF-beta signalling;

KM TGF-beta receptor type II; arterial remodelling; bone formation;

KM cartilage formation; osteogenesis imperfecta; Bethlem myopathy;

KM dystrophic epidermolysis bullosa; negative remodelling; wound healing;

KM arterial stenosis; fibrosis; calcification; transplant;

KM heart valve transplant; osteopathic; antihypertensive; vulnerable.

OS Homo sapiens.

XX US2002161211-A1.

XX 31-OCT-2002.

XX 19-OCT-2001; 2001US-00045992.

XX 19-OCT-2000; 2000US-00692081.

XX

PA (LIND) LINDER V.

PA (FRIE) FRIESEL R E.

XX Linder V, Friesel RE;

DR WPI; 2003-238238/23.

DR N-PSDB; ABX11341.

XX

PS Claim 3; Fig 4B; 81pp; English.

XX

CC The invention discloses an isolated nucleic acid encoding a mammalian

CC adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL

CC and/or adventitia induced bone expressed molecule, AIBE). REMODELIN

CC expression was induced by transforming growth factor beta (TGF-beta),

CC which is important because proliferative events occurring in the

CC adventitia contribute to vascular remodelling and restenosis in response

CC to vascular injury and TGF-beta has been shown to be a factor involved in

CC this. Also disclosed is an antibody raised against REMODELIN and methods

CC for treating a disease mediated by abnormal expression of a REMODELIN in

CC a human, for identifying a compound that affects or reduces expression of

CC REMODELIN in a cell (e.g. antisense therapy), for identifying a compound

CC that affects TGF-beta signalling and for increasing or reducing REMODELIN

CC expression in a mammal, comprising administering a REMODELIN expression

CC increasing or reducing TGF-beta to the mammal, thereby increasing

CC REMODELIN expression or inhibiting signalling via the TGF-beta receptor

CC type II and reducing expression of REMODELIN in the mammal. The methods

CC and compositions of the present invention are useful for mediating

CC arterial remodelling, formation of bone and cartilage and the diagnosis

CC and treatment of disorders associated with aberrant expression of

CC REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis

CC bullosa, Bethlem myopathy, negative remodelling, wound healing, arterial

CC stenosis, vessel injury, fibrosis and calcification of a transplant,

CC preferably a heart valve transplant. The sequence presented is the human

CC REMODELIN protein

XX

XX Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 6; Length 243;

Best Local Similarity 100.0%; Pred. No. 9e-121; Mismatches 0; Gaps 0;

Matches 243; Conservative 0; Indels 0; Gaps 0;

QY 1 MRPGPAPSPORLRLGLLLLLQLPAPSSASSEIPKQKQAKLRQREVVDLYNGMCIQGPA 60

1 MRPGPAPSPORLRLGLLLLLQLPAPSSASSEIPKQKQAKLRQREVVDLYNGMCIQGPA 60

QY 61 GVPGRDGSPGANGIPETPGIPGRDGFKEGECLERESFESWTPTYKQCSWSLNYGIDL 120

61 GVPGRDGSPGANGIPETPGIPGRDGFKEGECLERESFESWTPTYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRYTFENGACSGPLPIEAIITYLQ 180

121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRYTFENGACSGPLPIEAIITYLQ 180

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240

181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240

QY 241 LPK 243

241 LPK 243

DB 241 LPK 243

RESULT 8

ABU56607

ID ABU56607 standard; protein; 243 AA.

XX ABU56607;

XX

DT	02-APR-2003	(first entry)
DE	Lung cancer-associated polypeptide #200.	
XX		
XX	Lung cancer-associated polypeptide; cytostatic; emphysema;	
XX	antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;	
KW	small cell lung cancer; benign lesion; precancerous lesion; bronchitis;	
KW	chronic obstructive pulmonary disease; hypersensitivity pneumonitis;	
KW	interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.	
XX		
OS	Unidentified.	
XX		
PN	MO200286443-A2.	
XX		
PD	31-OCT-2002.	
XX		
PF	18-APR-2002; 2002WO-US012476.	
XX		
PR	18-APR-2001; 2001US-0284470P.	
PR	10-MAY-2001; 2001US-0290492P.	
PR	09-NOV-2001; 2001US-039245P.	
PR	13-NOV-2001; 2001US-035066P.	
PR	29-NOV-2001; 2001US-0334370P.	
XX	12-APR-2002; 2002US-0372246P.	
PA	(EOSB-) EOS BIOTECHNOLOGY INC.	
XX		
PI	Aziz N, Murray R;	
XX		
DR	WPI: 2003-093161/08.	
XX	N-PSDB; ABX76336.	
PT		
PT	Detecting a lung cancer-associated transcript in a cell from a patient	
PT	for treating lung cancer, by contacting a biological sample from the	
PT	patient with a polynucleotide that exhibits increased or decreased	
XX	expression in lung cancer.	
XX		
PS	Claim 27; Page 337; 453pp; English.	
XX		
XX	The invention relates to a method for detecting a lung cancer-associated	
CC	transcript in a cell from a patient, comprising contacting a biological	
CC	sample from the patient with a polynucleotide that selectively hybridises	
CC	to a sequence that is at least 80 % identical to a gene that exhibits	
CC	increased or decreased expression in lung cancer samples, lung cancer-	
CC	associated polynucleotides and polypeptides are used for identifying a	
CC	compound that modulates a lung cancer-associated polypeptide, for	
CC	inhibiting proliferation of a lung cancer-associated cell to treat lung	
CC	cancer in a patient and for treating a mammal having lung cancer by	
CC	administering a modulatory compound identified. The methods are useful	
CC	for treating lung cancer, such as small cell lung cancer, non-small cell	
CC	lung cancer or other benign or precancerous lesions, e.g. atelectasis,	
CC	emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis and	
CC	hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and	
CC	bronchiectasis. The genes, polynucleotides and polypeptides are useful	
CC	for diagnostic purposes and as targets for screening for therapeutic	
CC	compounds that modulate lung cancer, such as antibodies. Sequences	
CC	ABU56408-ABU56745 represent lung cancer-associated polypeptides of the	
CC	invention	
XX		
XX		
XX	Sequence 243 AA;	
XX		
Query Match	100.0%; Score 1303; DB 6; Length 243;	
Best Local Similarity	100.0%; Pred. No. 9e-121;	
Matches 243; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
DB	1 MRPGPAPSPORLRLGLLLILLIQLPAPSSASRIPIKQKQAKOLRQREVDVLYNMCILQGP 60	
DB	1 MRPGPAPSPQRLRLGLLLILLIQLPAPSSASRIPIKQKQAKQLRQREVDVLYNMCILQGP 60	
QY	61 GYPRDGSFGANGICPETPGIPGRDGGKGEKCECLRESFESWTPNPKQCSWSLNTGIDT 120	
QY	61 GYPRDGSFGANGICPETPGIPGRDGGKGEKCECLRESFESWTPNPKQCSWSLNTGIDT 120	
DB	61 GYPRDGSFGANGICPETPGIPGRDGGKGEKCECLRESFESWTPNPKQCSWSLNTGIDT 120	

Query	Match	Score	DB	Length	243
Query Match	100.0%	Score 1303	DB 7	Length 243	
Best Local Similarity	100.0%	Pred. No. 9e-121			

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAAPORLRLGLLLLLLLQLPAPSSASEIPKQKQAKQRLRREVVLDLYNMGCLQGPA 60  
 DB 1 MRPGPAAPORLRLGLLLLLLLQLPAPSSASEIPKQKQAKQRLRREVVLDLYNMGCLQGPA 60  
 QY 61 GVPGRDSSPGANGIPGTPTGIPGRDGFKEGKECELRSEFSESWTPNPKQCSWSSLNVTGIDL 120  
 DB 61 GVPGRDSSPGANGIPGTPTGIPGRDGFKEGKECELRSEFSESWTPNPKQCSWSSLNVTGIDL 120  
 QY 121 GKIAECTFTKMSNSALRVLFSSGLRLKCRNACCCQRYFTFNGAECGSLPIEAIITLDDQ 180  
 DB 121 GKIAECTFTKMSNSALRVLFSSGLRLKCRNACCCQRYFTFNGAECGSLPIEAIITLDDQ 180  
 QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITIEE 240  
 DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITIEE 240  
 QY 241 LPK 243  
 DB 241 LPK 243

RESULT 10  
 ID ADN39855 standard; protein; 243 AA.  
 AC ADN39855;  
 AC ADN39855;  
 DT 17-JUN-2004 (first entry)  
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C225.

XX Human; differential expression; cancer; angiogenic disorder;  
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
 KW inflammatory disease; autoimmune disease;  
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;  
 KW detection; diagnosis; prognosis; drug screening; drug targeting;  
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
 KW vulnery; gene therapy; vaccine.

XX Homo sapiens.  
 OS  
 PN MO2003042661-A2.  
 PD 22-MAY-2003.  
 PF 13-NOV-2002; 2002WO-US036810.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 21-NOV-2001; 2001US-0332464P.  
 PR 29-NOV-2001; 2001US-0334333P.  
 PR 03-DEC-2001; 2001US-0335394P.  
 PR 14-DEC-2001; 2001US-0340376P.  
 PR 08-JAN-2002; 2002US-0347211P.  
 PR 10-JAN-2002; 2002US-0347349P.  
 PR 08-FEB-2002; 2002US-0355250P.  
 PR 13-FEB-2002; 2002US-0356714P.  
 PR 20-FEB-2002; 2002US-0359077P.  
 PR 29-MAR-2002; 2002US-0368809P.  
 PR 04-APR-2002; 2002US-0370110P.  
 PR 12-APR-2002; 2002US-0372246P.  
 PR 05-JUN-2002; 2002US-0386614P.  
 PR 16-JUL-2002; 2002US-0396839P.  
 PR 22-JUL-2002; 2002US-0397845P.  
 PR 09-SEP-2002; 2002US-0409450P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 PA  
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;  
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;  
 XX

DR WPI: 2003-468649/44.  
 DR N-PSDB; ADN398537.  
 PT Determining the presence or absence of a pathological cell in a patient,  
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting  
 PT a nucleic acid in a biological sample.  
 PS Claim 12, SEQ ID NO C225; 1385BP; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
 CC whose expression is upregulated or downregulated in specific cancers or  
 CC other diseases such as angiogenic or fibrotic disorders, and to methods  
 CC of determining the presence or absence of a pathological cell in a  
 CC patient by detecting a nucleic acid at least 80% identical to those of  
 CC the invention or by detecting a polypeptide of the invention. The  
 CC invention also relates to expression vectors and host cells comprising a  
 CC nucleic acid of the invention; antibodies which specifically bind a  
 CC polypeptide of the invention; use of such antibodies for drug targeting;  
 CC and methods of screening for modulators of activity or expression of the  
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,  
 CC antibodies and methods are useful for diagnosing, prognosing and treating  
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,  
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
 CC neovascularisation syndromes, scarring and uterine fibroids. They may  
 CC also be useful in wound healing and in contraception. The present  
 CC sequence represents a polypeptide of the invention.

XX Sequence 243 AA;  
 SQ

Query Match 100.0%; Score 1303; DB 7; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 9e-121;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAAPORLRLGLLLLLLLQLPAPSSASEIPKQKQAKQRLRREVVLDLYNMGCLQGPA 60  
 DB 1 MRPGPAAPORLRLGLLLLLLLQLPAPSSASEIPKQKQAKQRLRREVVLDLYNMGCLQGPA 60  
 QY 61 GVPGRDSSPGANGIPGTPTGIPGRDGFKEGKECELRSEFSESWTPNPKQCSWSSLNVTGIDL 120  
 DB 61 GVPGRDSSPGANGIPGTPTGIPGRDGFKEGKECELRSEFSESWTPNPKQCSWSSLNVTGIDL 120  
 QY 121 GKIAECTFTKMSNSALRVLFSSGLRLKCRNACCCQRYFTFNGAECGSLPIEAIITLDDQ 180  
 DB 121 GKIAECTFTKMSNSALRVLFSSGLRLKCRNACCCQRYFTFNGAECGSLPIEAIITLDDQ 180  
 QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITIEE 240  
 DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITIEE 240  
 QY 241 LPK 243  
 DB 241 LPK 243

RESULT 11  
 ID ADN38732 standard; protein; 243 AA.  
 AC ADN38732;  
 AC ADN38732;  
 DT 17-JUN-2004 (first entry)  
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:50.

XX Human; differential expression; cancer; angiogenic disorder;  
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
 KW inflammatory disease; autoimmune disease;  
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;  
 KW detection; diagnosis; prognosis; drug screening; drug targeting;  
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
 KW vulnery; gene therapy; vaccine.

XX Homo sapiens.  
 OS

```

XX  MO20030426661-A2.
PN
XX
PD
XX
PF
XX
13-NOV-2002; 2002WO-US036810.
XX
13-NOV-2001; 2001US-0350666P.
PR
21-NOV-2001; 2001US-0332464P.
PR
29-NOV-2001; 2001US-0334393P.
PR
03-DEC-2001; 2001US-0335394P.
PR
14-DEC-2001; 2001US-0340376P.
PR
08-JAN-2002; 2002US-0347211P.
PR
10-JAN-2002; 2002US-0347349P.
PR
08-FEB-2002; 2002US-0355250P.
PR
13-FEB-2002; 2002US-0356714P.
PR
20-FEB-2002; 2002US-0359077P.
PR
29-MAR-2002; 2002US-036809P.
PR
04-APR-2002; 2002US-0370110P.
PR
12-APR-2002; 2002US-0372246P.
PR
05-JUN-2002; 2002US-038614P.
PR
16-JUL-2002; 2002US-0396839P.
PR
22-JUL-2002; 2002US-039775P.
PR
22-JUL-2002; 2002US-0397845P.
PR
09-SEP-2002; 2002US-0409450P.
XX
XX
PA
(EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI
Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;
PI
Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;
XX
XX
WPI; 2003-468649/44.
DR
N-PSDB; ADN38731.
XX
XX
PT
Determining the presence or absence of a pathological cell in a patient,
PT
useful for diagnosing, prognosing or treating cancer, comprises detecting
XX
a nucleic acid in a biological sample.
XX
PS
Claim 12; SEQ ID NO 50; 1385pp; English.
XX
XX
CC
The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC
whose expression is upregulated or downregulated in specific cancers or
CC
other diseases such as angiogenic or fibrotic disorders, and to methods
CC
of determining the presence or absence of a pathological cell in a
CC
patient by detecting a nucleic acid at least 80% identical to those of
CC
the invention or by detecting a polypeptide of the invention. The
CC
invention also relates to expression vectors and host cells comprising a
CC
nucleic acid of the invention; antibodies which specifically bind a
CC
polypeptide of the invention; use of such antibodies for drug targeting;
CC
and methods of screening for modulators of activity or expression of the
CC
polypeptides and nucleic acids. The nucleic acid, polypeptides,
CC
antibodies and methods are useful for diagnosing, prognosing and treating
CC
cancer and other conditions such as psoriasis, ischemia, heart disease,
CC
atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC
neovascularization syndromes, scarring and uterine fibroids. They may
CC
also be useful in wound healing and in contraception. The present
CC
sequence represents a polypeptide of the invention.
XX
XX
SQ
Sequence 243 AA;
XX
XX
Query Match 100.0%; Score 1303; DB 7; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY
1 MRQGPAAAPQRLRGILLILLILLPAPSSASEIPKQKQKQRLQREVVLDLYNMCLOQPA 60
DB
1 MRQGPAAAPQRLRGILLILLILLPAPSSASEIPKQKQKQRLQREVVLDLYNMCLOQPA 60
QY
61 GVGGRDGPSPANGIPGPGIPGRDGFKEGEGECLARSFESWPNPKQCSWSSLYNGID 120
DB
61 GVGGRDGPSPANGIPGPGIPGRDGFKEGEGECLARSFESWPNPKQCSWSSLYNGID 120
QY
121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYTFNGAGCSGPLPBAIYLDQ 180

```

```

DB
121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYTFNGAGCSGPLPBAIYLDQ 180
QY
181 GSPENNSTINIRHTSSVEGLCEGIGAGLVDAIWMGTCSDPKPGDASTGMSVSRITIEE 240
DB
181 GSPENNSTINIRHTSSVEGLCEGIGAGLVDAIWMGTCSDPKPGDASTGMSVSRITIEE 240
QY
241 LPK 243
DB
241 LPK 243
RESULT 12
ADL70254
ID ADL70254 standard; protein; 243 AA.
XX
AC ADL70254;
XX
DT 20-MAY-2004 (first entry)
XX
DE LBFJ301 polypeptide, associated with stomach cancer.
XX
KW LBFJ301; human; stomach cancer; diagnosis; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004016636-A1.
XX
PD 26-FEB-2004.
XX
PF 14-AUG-2003; 2003WO-KR001653.
XX
PR 14-AUG-2002; 2002US-0402904P.
PR 20-AUG-2002; 2002US-0404408P.
PR 23-AUG-2002; 2002US-0405304P.
PR 28-OCT-2002; 2002US-0421582P.
XX
PA (GLDS ) LG LIFE SCI LTD.
XX
PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Song SY;
XX
WPI; 2004-203757/19.
DR
N-PSDB; ADL70253.
XX
PT New nucleic acid molecule associated with stomach cancer, useful for
PT
preparing a composition for diagnosing or treating cancers such as
XX
stomach cancer.
XX
PS
Claim 24; SEQ ID NO 2; 146pp; English.
XX
XX
CC
The present sequence is that of the polypeptide encoded by cDNA clone
CC
AD12. This clone corresponds to an mRNA species that is differentially
CC
expressed in cancerous stomach tissue compared to healthy stomach tissue.
CC
Clones AD12 and CH4 ADL70255 are splice variants of a gene designated
CC
LBFJ301. Expression of this gene is significantly up-regulated (13.75-
CC
fold) in gastric carcinoma samples compared to samples from healthy
CC
stomach tissue. Up-regulation of LBFJ301 may therefore be diagnostic for
CC
stomach cancer. The predicted proteins encoded by AD12 and CH4 ADL70256
CC
are identical for the first 124 amino acids, while the last 13 amino
CC
acids of the CH4 protein are unique. Termination of the CH4 protein is
CC
produced by a 45 bp insertion which introduces a stop codon in the open
CC
reading frame. The proteins are weakly similar to the chymotrypsin serine
CC
protease family signature and the NUDIX hydrolase family signature. The
CC
nucleic acids and proteins of the invention can be used as diagnostic
CC
agents or markers to detect stomach cancer or to monitor the progression
CC
of stomach cancer in a sample. The proteins can be used to identify
CC
agents which modulate protein level or activity, to identify binding
CC
partners, to raise antibodies, as therapeutic targets, and as diagnostic
CC
agents or markers of stomach cancer and other hyperplastic diseases. A
CC
claimed method for diagnosing a disease state in a subject comprises
CC
determining the level of expression of the nucleic acid, where the
CC
disease state is stomach cancer, advanced gastric cancer or a malignant
CC
neoplasm occurring in soft tissue, bone, breast, cervix, colon,

```

CC endometrium, oesophagus, kidney, latex, liver, lung, omentum, ovary,  
 CC pancreas, rectum, thyroid, myometrium, prostate, skin, small intestine,  
 CC bladder, spleen or stomach.  
 XX  
 SQ Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 8; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 9e-121;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRLRGLLLLLQLPAPSSASEIPKQKQALRQREVVDLYNMCLOGPA 60  
 DB 1 MRPGPASPQRLRGLLLLLQLPAPSSASEIPKQKQALRQREVVDLYNMCLOGPA 60  
 QY 61 GVPGRDSSPGANGIPGTPIGIPGRDGFKEGCECLRESFEESWTPNYKQCSWSLNYGIDL 120  
 DB 61 GVPGRDSSPGANGIPGTPIGIPGRDGFKEGCECLRESFEESWTPNYKQCSWSLNYGIDL 120  
 QY 121 GKIAECTFTKMRNSALRVLFSSGLRLKCRNACCQRYFTFNGAECGSLPIEAIITYLDQ 180  
 DB 121 GKIAECTFTKMRNSALRVLFSSGLRLKCRNACCQRYFTFNGAECGSLPIEAIITYLDQ 180  
 QY 181 GSEPMNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDYPKGDASTGMNSVSRITIEE 240  
 DB 181 GSEPMNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDYPKGDASTGMNSVSRITIEE 240  
 QY 241 LPK 243  
 DB 241 LPK 243

## RESULT 13

AAB08856  
 ID AAB08856 standard; protein; 278 AA.

XX AC AAB08856;  
 XX DT 15-JAN-2001 (first entry)  
 XX DE Amino acid sequence of a human secretory protein.  
 XX KM Human; secretory protein; HSECP; cancer; gastrointestinal disorder;  
 XX KM inflammation; cardiovascular disorder; neurological disorder.  
 XX OS Homo sapiens.  
 XX FH Key  
 XX FH Peptide 1..65  
 FT /note= "signal sequence"  
 FT Modified-site 7  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 64  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 132  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 179  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 221  
 FT /note= "potential glycosylation site"  
 FT Modified-site 230  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 252  
 FT /note= "potential phosphorylation site"  
 XX PN W0200052151-A2.  
 XX PD 08-SEP-2000.  
 XX PF 03-MAR-2000; 2000MO-US005621.  
 XX PR 05-MAR-1999; 99US-0123117P.  
 XX PA (INCY-) INCYTE PHARM INC.

XX PI Tang YT, Lal P, Baughn MR, Yue H, Au-Young J, Lu DM, Azimzai Y;  
 XX WPI; 2000-579282/54.  
 DR N-PSDB; AAA75110.  
 XX

PT Twenty two human secretory proteins for diagnosing, treating and  
 PT preventing cancer, inflammation, and gastrointestinal, cardiovascular and  
 PT neurological disorders.  
 XX

PS Claim 1; Page 83; 107pp; English.

CC The present sequence represents a human secretory protein, designated  
 CC HSECP-1. The specification also describes HSECP-2 to HSECP-22. The  
 CC proteins are useful for diagnosing, treating and preventing cancer,  
 CC inflammation, and gastrointestinal, cardiovascular and neurological  
 CC disorders. The proteins may also be used to identify agonists,  
 CC antagonists, and inhibitors. The polynucleotides may be used for  
 CC producing the protein recombinantly, and as a source of probes and  
 CC primers for isolating and identifying related sequences

SQ Sequence 278 AA;

Query Match 100.0%; Score 1303; DB 3; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-120;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRLRGLLLLLQLPAPSSASEIPKQKQALRQREVVDLYNMCLOGPA 60  
 DB 36 MRPGPASPQRLRGLLLLLQLPAPSSASEIPKQKQALRQREVVDLYNMCLOGPA 95  
 QY 61 GVPGRDSSPGANGIPGTPIGIPGRDGFKEGCECLRESFEESWTPNYKQCSWSLNYGIDL 120  
 DB 61 GVPGRDSSPGANGIPGTPIGIPGRDGFKEGCECLRESFEESWTPNYKQCSWSLNYGIDL 120  
 QY 96 GVPGRDSSPGANGIPGTPIGIPGRDGFKEGCECLRESFEESWTPNYKQCSWSLNYGIDL 155  
 DB 96 GVPGRDSSPGANGIPGTPIGIPGRDGFKEGCECLRESFEESWTPNYKQCSWSLNYGIDL 155  
 QY 121 GKIAECTFTKMRNSALRVLFSSGLRLKCRNACCQRYFTFNGAECGSLPIEAIITYLDQ 180  
 DB 121 GKIAECTFTKMRNSALRVLFSSGLRLKCRNACCQRYFTFNGAECGSLPIEAIITYLDQ 180  
 QY 156 GKIAECTFTKMRNSALRVLFSSGLRLKCRNACCQRYFTFNGAECGSLPIEAIITYLDQ 215  
 DB 156 GKIAECTFTKMRNSALRVLFSSGLRLKCRNACCQRYFTFNGAECGSLPIEAIITYLDQ 215  
 QY 181 GSEPMNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDYPKGDASTGMNSVSRITIEE 240  
 DB 216 GSEPMNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDYPKGDASTGMNSVSRITIEE 275  
 QY 241 LPK 243  
 DB 241 LPK 243  
 DB 276 LPK 278

## RESULT 14

AAM25746  
 ID AAM25746 standard; protein; 278 AA.

XX AC AAM25746;  
 XX DT 16-OCT-2001 (first entry)  
 XX DE Human protein sequence SEQ ID NO:1261.  
 XX KM Human; cancer; HIV infection; human immunodeficiency virus;  
 KM antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
 KM antibacterial; endocrine; cardiac; central nervous system; virucide;  
 KM anti-HIV; fungicide; antimutagen; cardiovascular; antinaemic; anaemia;  
 KM antiaggregant; haemostatic; vulnerary; antitumor; osteopathic; eczema;  
 KM dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
 KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
 KM antiaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
 KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
 KM genetic disease; haematopoietic disorder; platelet disorder; asthma;  
 KM chromocytoma; osteoporosis; severe combined immunodeficiency;  
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;  
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 KM neurological disorder.

```
XX OS Homo sapiens.
XX XX
XX PN WO200153455-A2.
XX PD
XX PD 26-JUL-2001.
XX PF
XX PF 22-DEC-2000; 2000MO-US035017.
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-0052317.
XX PA
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR
XX DR MPI; 2001-457603/49.
XX DR N-PSDB; AAH99687.
XX PT
XX PT Isolated human polynucleotides encoding polypeptides, useful for the
XX PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX PS
XX PS Claim 20; Page 260; 1217pp; English.
XX CC
XX CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
XX CC AAM25963. The proteins can have activities based on the tissues and cells
XX CC they are expressed in, such as: antiinflammatory; antirheumatic;
XX CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;
XX CC central nervous system; virocidic; anti-HIV; fungicidal; antimutagen;
XX CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerrary;
XX CC antitumor; osteopathic; dermatological; antiallergic; antistatic;
XX CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
XX CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX CC encoding them can be used in gene therapy, antisense therapy and vaccine
XX CC production. The proteins and polynucleotides are useful for screening for
XX CC agonists or antagonists of a protein and for the treatment and diagnosis
XX CC of disorders associated with the activity of a protein e.g. inflammation,
XX CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
XX CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
XX CC osteoporosis, severe combined immunodeficiency, eczema, allergic
XX CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
XX CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
XX CC neurological disorders
XX SQ
XX SQ Sequence 278 AA;
XX
XX Query Match 100.0%; Score 1303; DB 4; Length 278;
XX Best Local Similarity 100.0%; Pred. No. 1,1e-120;
XX Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MRQGPAPASPORLRGLLLLLLLQLPAPSSASERPKGKQALRQREVDVLYNMCLOGPA 60
XX DB 36 MRQGPAPASQRLRGLLLLLLLQLPAPSSASERPKGKQALRQREVDVLYNMCLOGPA 95
XX QY 61 GVGGRDSSPGANGIPGTPTGIPGDDGRKGEKGECLRSFESSWTPNTYKQCSWSSLYATIDL 120
XX DB 96 GVGGRDSSPGANGIPGTPTGIPGDDGRKGEKGECLRSFESSWTPNTYKQCSWSSLYATIDL 155
XX QY 121 GKLAECFTKMRNSALRVLFSGSLRLKCNACCORRYFPFNABCSGPIPTAIIYLDQ 180
XX DB 156 GKLAECFTKMRNSALRVLFSGSLRLKCNACCORRYFPFNABCSGPIPTAIIYLDQ 215
XX QY 181 GSPFNASTINIRHTSSVEGLCEGIGAGLVVAITWGTCSDPYKGDASTGWNVSRIIIEE 240
XX DB 216 GSPFNASTINIRHTSSVEGLCEGIGAGLVVAITWGTCSDPYKGDASTGWNVSRIIIEE 275
XX QY 241 LPK 243
XX |||
XX DB 276 LPK 278
```

```
RESULT 15
ID ABG96338
ID ABG96338 standard; protein; 278 AA.
XX
XX AC ABG96338;
XX XX
XX DT 11-DEC-2002 (first entry)
XX
XX DE Human ovarian cancer marker M138.
XX
XX KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
XX KW central nervous system disorder; bacterial meningitis; viral meningitis;
XX KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
XX KW brain herniation; inflammation; encephalitis; testicular disorder;
XX KW nonbuberculous granulomatous orchitis; connective tissue disorder;
XX KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
XX KW histological type; carcinogenic; ovarian cancer marker.
XX
XX OS Homo sapiens.
XX
XX PN WO200271928-A2.
XX
XX PD 19-SEP-2002.
XX
XX PF 14-MAR-2002; 2002MO-US007826.
XX
XX PR 14-MAR-2001; 2001US-0276025P.
XX PR 14-MAR-2001; 2001US-0276026P.
XX PR 10-AUG-2001; 2001US-0311732P.
XX PR 19-SEP-2001; 2001US-0323580P.
XX PR 26-SEP-2001; 2001US-0324967P.
XX PR 26-SEP-2001; 2001US-0325102P.
XX PR 26-SEP-2001; 2001US-0325149P.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX
XX PI Monahan JE, Gamavarrapu M, Hoerach S, Kamatkar S, Kovattis SG;
XX PI Meyers RE, Morrissey MP, Olandt RJ, Sen A, Vleebj PO, Mills GB;
XX PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
XX DR MPI; 2002-723277/78.
XX
XX DR N-PSDB; ABS76434.
XX
XX PT Assessing whether a patient is afflicted with ovarian cancer, useful in
XX PT assessing the stage or progression of the disease, comprises comparing
XX PT the expression level of a cancer marker in a sample from a patient and
XX PT from a non cancer patient.
XX
XX PS Disclosure; Page 249-250; 481pp; English.
XX
XX CC The present invention relates to a new method for assessing whether a
XX CC patient is afflicted with ovarian cancer. The method involves comparing
XX CC the expression level of a marker in a patient sample and the normal level
XX CC of expression of the marker in a control non-ovarian cancer sample, where
XX CC the marker is selected from 363 cancer markers described in the
XX CC specification. The method of the invention is useful in diagnosing or
XX CC characterising cancer, in detecting the presence of cancer as early as
XX CC possible, and the recurrence of ovarian cancer. The method may also be of
XX CC particular use with patients having an enhanced risk of developing
XX CC ovarian cancer (e.g. patients having a familial history of ovarian
XX CC cancer). The cancer markers may be used in the management and treatment
XX CC of e.g. brain and central nervous system disorders (e.g. bacterial and
XX CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
XX CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
XX CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
XX CC testicular disorders (e.g. nonbuberculous granulomatous orchitis),
XX CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
XX CC disease or atherosclerosis). The compositions and methods may also be
XX CC used in assessing the histological type of neoplasm associated with
XX CC ovarian cancer, monitoring the progression of ovarian cancer, determining
XX CC whether ovarian cancer has metastasized or is likely to metastasize,
XX CC selecting a composition for inhibiting ovarian cancer, assessing the
```





This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:56 ; Search time 24.4588 Seconds  
(without alignments)  
741.643 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303  
Sequence: 1 MRPGPAASPORLRLGLLL.....GDASTGMSVSRILIELEK 243

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	4 US-09-692-081-4	Sequence 4, Appl1
2	1300	99.8	243	4 US-09-489-847-205	Sequence 205, App
3	1294	99.3	243	4 US-09-834-759-514	Sequence 514, App
4	1294	99.3	278	4 US-09-834-759-515	Sequence 515, App
5	1225	94.0	245	4 US-09-692-081-2	Sequence 2, Appl1
6	1225	94.0	277	4 US-09-692-081-5	Sequence 5, Appl1
7	1109.5	85.1	276	4 US-09-205-258-958	Sequence 58, App
8	1046	80.3	197	4 US-09-834-759-516	Sequence 516, App
9	1046	80.3	232	4 US-09-834-759-517	Sequence 517, App
10	356	27.3	66	4 US-09-205-258-962	Sequence 962, App
11	296	22.7	52	4 US-09-205-258-961	Sequence 961, App
12	270	20.7	51	4 US-09-205-258-963	Sequence 963, App
13	248	19.0	46	4 US-09-834-759-518	Sequence 518, App
14	220.5	16.9	52	4 US-09-205-258-960	Sequence 960, App
15	196	15.0	93	4 US-09-489-847-240	Sequence 240, App
16	196	15.0	93	4 US-09-489-847-362	Sequence 362, App
17	141	10.8	26	4 US-09-834-759-519	Sequence 519, App
18	131.5	10.1	246	2 US-08-463-911-4	Sequence 4, Appl1
19	127.5	9.8	1057	3 US-08-931-820-1	Sequence 1, Appl1
20	127.5	9.8	1461	4 US-09-585-887-9	Sequence 9, Appl1
21	127.5	9.8	1461	4 US-09-289-578-9	Sequence 21, Appl1
22	127.5	9.8	1464	4 US-09-311-347C-21	Sequence 2, Appl1
23	125	9.6	313	4 US-09-949-016-9265	Sequence 9265, App
24	125	9.6	1670	4 US-09-949-016-5883	Sequence 5883, App
25	124.5	9.6	128	3 US-09-227-357-190	Sequence 190, App
26	124.5	9.6	1516	4 US-09-949-016-8209	Sequence 8209, App
27	124	9.5	623	3 US-09-029-348-3	Sequence 3, Appl1

28	124	9.5	626	3 US-09-029-348-2	Sequence 2, Appl1
29	122.5	9.4	1268	4 US-09-949-016-7487	Sequence 7487, App
30	121.5	9.3	492	4 US-08-468-996-12	Sequence 12, App
31	121	9.3	96	4 US-09-513-999C-4206	Sequence 4206, App
32	121	9.3	208	4 US-09-471-276-823	Sequence 823, App
33	121	9.3	208	4 US-09-800-729-151	Sequence 151, App
34	120	9.2	245	4 US-09-552-225A-4	Sequence 4, Appl1
35	120	9.2	245	4 US-09-552-204A-4	Sequence 4, Appl1
36	119.5	9.2	310	3 US-09-219-849-47	Sequence 47, Appl1
37	119.5	9.2	595	3 US-09-219-849-48	Sequence 48, Appl1
38	119.5	9.2	595	3 US-09-219-849-50	Sequence 50, Appl1
39	119.5	9.2	684	3 US-08-555-663-12	Sequence 12, Appl1
40	119.5	9.2	684	3 US-09-073-663-12	Sequence 12, Appl1
41	119.5	9.2	755	4 US-09-919-497-57	Sequence 57, Appl1
42	119.5	9.2	822	3 US-09-219-849-49	Sequence 49, Appl1
43	119	9.1	258	4 US-09-976-594-815	Sequence 815, App
44	119	9.1	1806	4 US-09-919-497-56	Sequence 56, Appl1
45	118.5	9.1	1694	1 US-08-494-168-2	Sequence 2, Appl1

# ALIGNMENTS

```

RESULT 1
US-09-692-081-4
; Sequence 4, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-692-081-4

Query Match      100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 6.5e-125;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKLROREVDLYNGMCIQGA 60
D      1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKLROREVDLYNGMCIQGA 60
QY      61 GVPGRDGSFGANGIPETPGIPGRDGFKGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120
D      61 GVPGRDGSFGANGIPETPGIPGRDGFKGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120
D      61 GVPGRDGSFGANGIPETPGIPGRDGFKGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120
QY      121 GGIABTFTPKMSNSALRYLFGSLRLKCRNACCQRYTFMGACSGSLPEATLYLDO 180
D      121 GGIABTFTPKMSNSALRYLFGSLRLKCRNACCQRYTFMGACSGSLPEATLYLDO 180
QY      181 GSPENNSTINIRTSVEGLCEGIGAGLVDAIWMGTCSDDYKPGDASTGMSVSRILIE 240
D      181 GSPENNSTINIRTSVEGLCEGIGAGLVDAIWMGTCSDDYKPGDASTGMSVSRILIE 240
QY      241 LPK 243
D      241 LPK 243

RESULT 2
US-09-489-847-205
; Sequence 205, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al

```

```

; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; EARLIER FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 205
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-205
```

```

Query Match          99.8%; Score 1300; DB 4; Length 243;
Best Local Similarity 99.6%; Pred. No. 1.3e-124;
Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MRPGPAASPORLRGLLLLLQLPAPSSASIEPKGKKAQLRQREVVLDLYNMGCLQGPA 60
DB      1 MRPGPAASPORLRGLLLLLQLPAPSSASIEPKGKKAQLRQREVVLDLYNMGCLQGPA 60
QY      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKEGECLEESFEESWTPTYKCCSMSSLYNGIDL 120
DB      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKEGECLEESFEESWTPTYKCCSMSSLYNGIDL 120
QY      121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSPLEAIYYIDQ 180
DB      121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSPLEAIYYIDQ 180
QY      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240
DB      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240
QY      241 LPK 243
DB      241 LPK 243
```

```

RESULT 3
US-09-834-759-514
; Sequence 514, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jjiang, Yugui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 514
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-514
```

```

Query Match          99.3%; Score 1294; DB 4; Length 243;
Best Local Similarity 99.6%; Pred. No. 5.4e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 MRPGPAASPORLRGLLLLLQLPAPSSASIEPKGKKAQLRQREVVLDLYNMGCLQGPA 60
DB      1 MRPGPAASPORLRGLLLLLQLPAPSSASIEPKGKKAQLRQREVVLDLYNMGCLQGPA 60
QY      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKEGECLEESFEESWTPTYKCCSMSSLYNGIDL 120
DB      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKEGECLEESFEESWTPTYKCCSMSSLYNGIDL 120
QY      121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSPLEAIYYIDQ 180
DB      121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSPLEAIYYIDQ 180
QY      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240
DB      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240
QY      241 LPK 243
DB      241 LPK 243
```

```

RESULT 4
US-09-834-759-515
; Sequence 515, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
```

```

; APPLICANT: Jjiang, Yugui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-515
```

```

Query Match          99.3%; Score 1294; DB 4; Length 278;
Best Local Similarity 99.6%; Pred. No. 6.5e-124;
Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 MRPGPAASPORLRGLLLLLQLPAPSSASIEPKGKKAQLRQREVVLDLYNMGCLQGPA 60
DB      36 MRPGPAASPORLRGLLLLLQLPAPSSASIEPKGKKAQLRQREVVLDLYNMGCLQGPA 95
QY      61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKEGECLEESFEESWTPTYKCCSMSSLYNGIDL 120
DB      96 GVPGRDGSFGANGIPGTGPIGRDGFKEGKEGECLEESFEESWTPTYKCCSMSSLYNGIDL 155
QY      121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSPLEAIYYIDQ 180
DB      156 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSPLEAIYYIDQ 215
QY      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240
DB      216 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 275
QY      241 LPK 243
```



Query Match	Similarity	85.1%	Score 1109.5	DB: 4	Length 276
Best local	Similarity	97.1%	Pred. No. 4,4e-105		
Matches	Conservative	3	Mismatches	2	Indels 1; Gaps 1
QY	34	PKGKQAKQLRREVVLDLYNGMCLQGPAVPGRGDGPANGIPETPGIPGRDGFGEKGE	93		
		:: ::			
Db	68	PRG-SKRRIRQREVVLDLYNGMCLQGPAVPGRGDGPANGIPETPGIPGRDGFGEKGE	126		
QY	94	LRSEFEESWTPMYKQCSWSLNYGIDLGIAETPTFKMSNSMLRVLFGSLRLKRRNC	153		
		LRSEFEESWTPMYKQCSWSLNYGIDLGIAETPTFKMSNSMLRVLFGSLRLKRRNC	186		
QY	154	CORWFFTEFGACSGPLPIEALITYLDQSGPENNSTINIRHTSSVEGLCGIGAGLVDAI	213		
Db	187	CORWFFTEFGACSGPLPIEALITYLDQSGPENNSTINIRHTSSVEGLCGIGAGLVDAI	246		
QY	214	WVGTCSDYPKGDASTGMSVSRITIIIEELPK	243		
		WVGTCSDYPKGDASTGMSVSRITIIIEELPK			
Db	247	WVGTCSDYPKGDASTGMSVSRITIIIEELPK	276		
RESULT 8 US-09-834-759-516 ; Sequence 516, Application US/09834759 ; Patent No. 6680197 ; GENERAL INFORMATION:					

Query Match	80.3%: Score 1046; DB 4; Length 197;
Best Local Similarity 99.5%; Pred. No. 8.5e-99;	
Matches 196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 1 MRPGAPASPORLRLGILLILLILQLPAPSSASEIPKQKQKQLRQREVVDLYNGCLOGPA 60	
DB 1 MRPGAPASPORLRLGILLILLILQLPAPSSASEIPKQKQKQLRQREVVDLYNGCLOGPA 60	
QY 61 GVPGRDPSPGANGIPGTPGIPGRDGFGEKGECELRSEFBSWTPNYKQCSWSSLYNGIDL 120	
DB 61 GVPGRDPSPGANVYPGPFGIPGRGFGKEKGECELRSEFBSWTPNYKQCSWSSLYNGIDL 120	
QY 121 GKIECEFTTKRSASARLVFFSGSLRLKCRNACCORNYFTFNNGACSGPLTEAIIYLDQ 180	
DB 121 GKIECEFTTKRSASARLVFFSGSLRLKCRNACCORNYFTFNNGACSGPLTEAIIYLDQ 180	
QY 181 GSPENNSTINIHRTSSV 197	
DB 181 GSPENNSTINIHRTSSV 197	
RESULT 9	
US-09-834-759-517	
; Sequence 517, Application US/09834759	
; Patent No. 6680197	
GENERAL INFORMATION:	
APPLICANT: Jiang, Yugu	
APPLICANT: Dillon, Davin C.	
APPLICANT: Mitcham, Jennifer L.	
APPLICANT: Xu, Jiangchun	
APPLICANT: Harlocker, Susan L.	
APPLICANT: Hepler, William T.	
APPLICANT: Henderson, Robert A.	
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND	
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER	
FILE REFERENCE: 210121.470C9	
CURRENT APPLICATION NUMBER: US/09/834,759	
CURRENT FILING DATE: 2001-04-13	
NUMBER OF SEQ ID NOS: 547	
SOFTWARE: FastSeq for Windows Version 3.0	
SEQ ID NO 517	
LENGTH: 232	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-09-834-759-517	
Query Match 80.3%; Score 1046; DB 4; Length 232;	
Best Local Similarity 99.5%; Pred. No. 1.1e-98;	
Matches 196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 1 MRPGAPASPORLRLGILLILLILQLPAPSSASEIPKQKQKQLRQREVVDLYNGCLOGPA 60	
DB 36 MRPGAPASPORLRLGILLILLILQLPAPSSASEIPKQKQKQLRQREVVDLYNGCLOGPA 95	

```
QY 61 GYRGGRGSPGANGIPGTPIGPDGFRGKGECELRSEFSESWTPNTYKQCSWSLNGIDL 120
DB 96 GYVGRDGSPPANVPIGTPGIGRDGFRGKGECELRSEFSESWTPNTYKQCSWSLNGIDL 155
QY 121 GKIAECTFTMRNSALRYLFGSGLRLKCRNACCCORWYTFNGAECGSLPIEAIYYLDQ 180
DB 156 GKIAECTFTMRNSALRYLFGSGLRLKCRNACCCORWYTFNGAECGSLPIEAIYYLDQ 215
QY 181 GSPEMNSTINIRHTSSV 197
DB 216 GSPEMNSTINIRHTSSV 232

RESULT 10
US-09-205-258-962
; Sequence 962, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
```

```
EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 962
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-962

Query Match 27.3%; Score 356; DB 4; Length 66;
Best Local Similarity 100.0%; Pred. No. 5.5e-29;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 FTMRNSALRYLFGSGLRLKCRNACCCORWYTFNGAECGSLPIEAIYYLDQSPENMS 187
DB 1 FTMRNSALRYLFGSGLRLKCRNACCCORWYTFNGAECGSLPIEAIYYLDQSPENMS 60

QY 188 TINHR 193
DB 61 TINHR 66

RESULT 11
US-09-205-258-961
; Sequence 961, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-961

Query Match          22.7%: Score 296; DB 4; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.3e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 77 TPGIGRDFGKGEKCECLRSFEESSWTPNYKQCSWSSLNYGIDGKIACECTF 128
Db 1 TPGIRGDFGKGEKCECLRSFEESSWTPNYKQCSWSSLNYGIDGKIACECTF 52

RESULT 12
US-09-205-258-963
; Sequence 963, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06

```



```

; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 963
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-963
```

```

Query Match          20.7%; Score 270; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 2,3e-20;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      193 RTSSVEGLCGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRRIIEELPK 243
Db      1 RTSSVEGLCGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRRIIEELPK 51
```

```

RESULT 13
US-09-834-759-518
; Sequence 518, Application US/09834759
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 518
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-518

Query Match          19.0%; Score 248; DB 4; Length 46;
```

```

Best Local Similarity 100.0%; Pred. No. 3.6e-18;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      198 EGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRRIIEELPK 243
Db      1 EGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRRIIEELPK 46
```

```

RESULT 14
US-09-205-258-960
; Sequence 960, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 960
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-960
```

```
Query Match      16.9%; Score 220.5; DB 4; Length 52;
Best Local Similarity 87.2%; Pred. No. 2.7e-15;
Matches 41; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
```

```
QY      34 PRGKQAKQLRQREVDLYNGMCLQSPAGVGRDGPANGIPGTPI 80
DB      7 PRG-SKRLIRQREVDLYNGMCLQSPAGVGRDGPANGIPGTPI 52
```

```
RESULT 15
US-09-489-847-240
; Sequence 240, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 240
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-240
```

```
Query Match      15.0%; Score 196; DB 4; Length 93;
```

Best Local Similarity 100.0%; Pred. No. 1.9e-12;  
Matches 40; Conservative 0; Mismatches 0; Indels 0;

```
QY      1 MRPGSPASPORTRLRGLLLLLLLQLPASPASSEIPKQKKA 40
DB      1 MRPGSPASPORTRLRGLLLLLLLQLPASPASSEIPKQKKA 40
```

Search completed: August 4, 2005, 04:23:53  
Job time : 25.4588 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:16 ; Search time 19.3765 Seconds

(without alignments)  
1206.653 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303  
Sequence: 1 MRQGPAPASQRLRGILL.....GDASTGMSVSRRIIEELK 243

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	131.5	10.1	246	2 S29328	complement subcomp
2	129	9.9	1752	2 A45407	collagen alpha 3(I)
3	128	9.8	289	2 T20177	hypothetical prote
4	127.5	9.8	1464	1 CGHUI5	collagen alpha 1(I)
5	126	9.7	327	2 T29031	hypothetical prote
6	125	9.7	754	2 A55267	collagen alpha 5(I)
7	125	9.6	888	2 S28791	collagen alpha 1(X)
8	125	9.6	1670	1 CGHUI3	collagen alpha 3(I)
9	124	9.5	1466	1 CGHUI7L	collagen alpha 1(I)
10	124	9.5	1763	2 S16366	collagen alpha 2(I)
11	123.5	9.5	311	2 T15268	hypothetical prote
12	123	9.4	304	2 T26185	hypothetical prote
13	123	9.4	304	2 T26184	hypothetical prote
14	123	9.4	1691	1 S22917	collagen alpha 5(I)
15	122.5	9.4	300	2 T24482	hypothetical prote
16	122.5	9.4	920	2 A45748	collagen alpha 1(V)
17	122.5	9.4	1549	2 T48103	collagen alpha 1(V)
18	122	9.4	2944	2 A54849	collagen alpha 1(V)
19	121	9.3	178	2 A39762	collagen alpha 1(X)
20	121	9.3	325	2 T18594	hypothetical prote
21	120.5	9.2	428	2 T24769	hypothetical prote
22	120	9.2	245	1 C1HUC	complement subcomp
23	120	9.2	358	2 T26281	hypothetical prote
24	119.5	9.2	298	2 T27644	hypothetical prote
25	119.5	9.2	310	2 T29731	hypothetical prote
26	119.5	9.2	458	2 T31631	hypothetical prote
27	119.5	9.2	671	1 CGRTIS	collagen alpha 1(I)
28	119.5	9.2	1453	1 S21626	collagen alpha 1(I)
29	119.5	9.2	1669	1 CGHUI4B	collagen alpha 1(I)

30	119	9.1	298	2 JCI448	collagen col-34 -
31	119	9.1	299	2 T29956	hypothetical prote
32	119	9.1	1042	1 CGCHIS	collagen alpha 1(I)
33	119	9.1	1049	1 CGBO7S	collagen alpha 1(I)
34	119	9.1	1806	1 CGHUIE	collagen alpha 1(X)
35	118.5	9.1	177	2 S37749	collagen alpha 2(X)
36	118.5	9.1	290	2 T24586	hypothetical prote
37	118.5	9.1	330	2 S46657	collagen alpha 1(X)
38	118.5	9.1	488	2 A27353	collagen alpha 1(I)
39	118.5	9.1	1691	1 CGHUI6B	collagen alpha 6(I)
40	118	9.1	1669	1 CGMS4B	collagen alpha 1(I)
41	118	9.1	1744	2 S40391	collagen alpha 1(I)
42	117.5	9.0	323	2 A61396	collagen alpha 1(I)
43	117.5	9.0	1486	1 B40333	collagen alpha 1(I)
44	117.5	9.0	1492	2 A40333	collagen alpha 1(I)
45	117	9.0	248	1 LNHUP1	pulmonary surfacta

#### ALIGNMENTS

##### RESULT 1

S29328 complement subcomponent C1q chain C - mouse

C/Species: Mus musculus (house mouse)

C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C/Accession: S29328

R/Petry, F.; Reid, K.B.M.; Loos, M.

Eur. J. Biochem. 209, 129-134, 1992

A/Title: Isolation, sequence analysis and characterization of cDNA clones coding for th

ecerebellin.

A/Reference number: S29328; MUID:93011118; PMID:1396691

A/Accession: S29328

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-246 <PEP>

A/Cross-references: UNIPROT:002105; EMBL:X66295; NID:950228; PTDN:CAA4693.1; PTD:95022

C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom

F;122-245/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 10.1%; Score 131.5; DB 2; Length 246;

Best Local Similarity 38.5%; Pred. No. 0.00026;

Matches 35; Conservative 5; Mismatches 32; Indels 19; Gaps 2;

QY	5	GPASFORLRLGLLLLLLPAPSSASERPKQKQAKLRQREVVDLYNKGCLQGPAGVPG	64
DB	4	GPSQCPQCGICLLFLFLALPLRSQAS-----ACGYGIPGMGMDGAGG	47
QY	65	ROGS---PGANGIPGPRGIPGRDGFGEKGE	92
DB	48	KDHGDLGQPKSPGIPAVGTGPRQKQGE	78

##### RESULT 2

A45407 collagen alpha 3(IV) chain - sea urchin (Strongylocentrotus purpuratus)

C/Species: Strongylocentrotus purpuratus (purple urchin)

C/Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C/Accession: A45407; A43903; A23940

R/Exposito, J.Y.; D'Alessio, M.; Di Liberto, M.; Ramirez, F.

J. Biol. Chem. 268, 5249-5254, 1993

A/Title: Complete primary structure of a sea urchin type IV collagen alpha chain and an

A/Reference number: A45407; MUID:93186842; PMID:844899

A/Accession: A45407

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: nucleic acid

A/Residues: 1-1752 <EXP>

A/Cross-references: UNIPROT:Q26312

A/Note: Sequence extracted from NCBI backbone (NCBIP:126841)

R/Wessel, G.M.; Etkin, M.; Benson, S.

Dev. Biol. 148, 261-272, 1991

A/Title: Primary mesenchyme cells of the sea urchin embryo require an autonomously produ

A/Reference number: A43903; MUID:92038439; PMID:1936564

A:Accession: A43903  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 'P', 633-1537, 'G' <WES>  
A:Cross-references: GB:S64572; NID:G238616; PIDN:AA80270.1; PID:G238617  
A:Note: sequence extracted from NCBI backbone (NCBIN:64572, NCBI:P:64573)  
R:Venkatesan, M.; De Pablo, F.; Vogeli, G.; Simpson, R.T.  
Proc. Natl. Acad. Sci. U.S.A. 83, 3351-3355, 1986  
A:Title: Structure and developmentally regulated expression of a Strongylocentrotus purpur  
A:Reference number: A23940; MUID:86205894; PMID:3458186  
A:Accession: A23940  
A:Molecule type: DNA  
A:Residues: 742-812 <VEN>  
A:Cross-references: EMBL:M13206  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix  
F:29-161/Domain: amino-terminal nonhelical, 7S <7SD>  
F:162-1523/Region: interrupted helical  
F:1524-1752/Domain: carboxyl-terminal nonhelical, NC1 <NC1>  
F:1534-1634/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F:1644-1748/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F:129/Modified site: allysine (Lys) #status predicted  
  
Query Match 9.9%; Score 129; DB 2; Length 1752;  
Best Local Similarity 22.9%; Pred. No. 0.004;  
Matches 47; Conservative 20; Mismatches 60; Indels 78; Gaps 9;  
  
QY 3 PGGPAAAPORLRLGLLLLLLLQLPAPSSASEIPKQKQALRQREVVDLYNMGCLQGPAGV 62  
1453 PGPPPRDR-----PGPG-----PG-----LTGDKGT 1477  
DB 63 PGRDPSRPGANGIPGTGPGRQSGFKGEGKCLRESF-----EESMT 103  
1478 PGVQGPFGVGVGPEPLKGBQCFRQCPGPGPGPKKGEGALPGSSGPFITRHSQT 1537  
QY 104 PNYKQC-----SW-----SSLNYGIDIGKIAECTFTMRSNSALRYLFGSGLT 147  
1538 TSPQCPQGTAKMWHYSILFLVQGNRGRGQDLGKRGSC-----LKRFTMPFLFC-NINN 1592  
DB 148 KCRNACCQGWYFTFGACSGGLPI 172  
1593 VCHVASRNDYSYWLSTTE---PWPM 1614  
QY 1593 VCHVASRNDYSYWLSTTE---PWPM 1614  
DB 1593 VCHVASRNDYSYWLSTTE---PWPM 1614  
  
RESULT 3  
T20177  
Hypothetical protein C53B4.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #ext\_change 09-Jul-2004  
C:Accession: T20177  
R:Berke, M.  
Submitted to the EMBL Data Library, December 1995  
A:Reference number: Z19233  
A:Accession: T20177  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-289 <WIL>  
A:Cross-references: UNIPROT:Q18799; EMBL:Z68215; PIDN:CAA92453.1; GSPDB:GN00022; CESP:CS  
A:Experimental source: Clone C53B4  
A:Genetic: C53B4.5  
A:Gene: CESP:C53B4.5  
A:Map position: 4  
  
Query Match 9.8%; Score 128; DB 2; Length 289;  
Best Local Similarity 35.7%; Pred. No. 0.00065;  
Matches 30; Conservative 5; Mismatches 19; Indels 30; Gaps 2;  
  
QY 25 PAPSASSEIPKQKQXQLRQREVVDLYNMGCLQGPAGVGRDSS-----P 69  
DB 203 PDPGSGEGRPGQPS-----RPGAGPQKQAGCGPBGKANGEPGP 247  
QY 70 GANGLPGRDGFPGKGEKGC 93

DB 248 GRDGPGRGPRGPRGRDGHPEKGV 271  
  
RESULT 4  
CGHUIS  
collagen alpha 1(I) chain precursor - human  
N:Alternate names: procollagen alpha 1(I) chain  
C:Species: Homo sapiens (man)  
C:Date: 12-Aug-1981 #sequence\_revision 04-Oct-1996 #ext\_change 09-Jul-2004  
C:Accession: 160114; S01143; A93335; I55254; A39943; I55237; A35233; S09400; B90567; S11  
5269; A29439; I53466; A02852; I37247  
R:D'Alessio, M.; Bernard, M.; Pretorius, P.J.; de Wet, W.; Ramirez, F.; Pretorius, P.J.  
Gene 67, 105-115, 1988  
A:Title: Complete nucleotide sequence of the region encompassing the first twenty-five e  
A:Reference number: 160114; MUID:86329734; PMID:284432  
A:Accession: 160114  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-369, 'L', 371-589 <DAL>  
A:Cross-references: UNIPROT:P02452; UNIPROT:Q14992; UNIPROT:Q16053; UNIPROT:Q13896; UNIP  
R:Tromp, G.; Kuivaniemi, H.; Stacey, A.; Shikata, H.; Baldwin, C.T.; Jaenisch, R.; Prock  
Biochem. J. 253, 919-922, 1988  
A:Title: Structure of a full-length cDNA clone for the prepro-alpha-1(I) chain of human  
A:Reference number: S01143; MUID:89025644; PMID:3178743  
A:Accession: S01143  
A:Molecule type: mRNA  
A:Residues: 1-472 <TRO>  
A:Cross-references: EMBL:X07884; NID:G30015; PIDN:CAA30731.1; PID:G30016; GB:M3646; NID  
A:Note: submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988  
R:Chu, M.L.; de Wet, W.; Bernard, M.; Ding, J.F.; Morabito, M.; Myers, J.; Williams, C.;  
Nature 310, 337-340, 1984  
A:Title: Human prolalpha1(I) collagen gene structure reveals evolutionary conservation of  
A:Reference number: A93335; MUID:84270697; PMID:646220  
A:Accession: A93335  
A:Molecule type: DNA  
A:Residues: 1-58, 'Q', 60-181 <CHU>  
A:Cross-references: EMBL:X00820; NID:G35657; PIDN:CAA2394.1; PID:G35658  
R:Rosow, C.M.S.; Vergeer, W.P.; du Plooy, S.J.; Bernard, M.P.; Ramirez, F.; de Wet, W.  
J. Biol. Chem. 262, 15151-15157, 1987  
A:Title: DNA sequences in the first intron of the human pro-alpha 1(I) collagen gene enh  
A:Reference number: I55254; MUID:88033098; PMID:2822714  
A:Accession: I55254  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-45 <ROS>  
A:Cross-references: GB:J02829; NID:G180387; PIDN:AA51993.1; PID:G180388  
R:Bornstein, P.; McKay, J.; Morishima, J.K.; Devarayalu, S.; Gelinas, R.E.  
Proc. Natl. Acad. Sci. U.S.A. 84, 8869-8873, 1987  
A:Title: Regulatory elements in the first intron contribute to transcriptional control o  
A:Reference number: A39943; MUID:88097389; PMID:3480516  
A:Accession: A39943  
A:Molecule type: DNA  
A:Residues: 1-34 <BOR>  
A:Cross-references: GB:J03559; NID:G180876; PIDN:AA52052.1; PID:G553238  
R:Chu, M.L.; de Wet, W.; Bernard, M.; Ramirez, F.  
J. Biol. Chem. 260, 2315-2320, 1985  
A:Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter s  
A:Reference number: I55237; MUID:85130970; PMID:2857713  
A:Accession: I55237  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-34 <CH2>  
A:Cross-references: GB:M0627; NID:G180383; PIDN:AA51992.1; PID:G553226  
R:Wirtz, W.K.; Keene, D.R.; Hori, H.; Glanville, R.W.; Steinmann, B.; Rao, V.H.; Hollist  
J. Biol. Chem. 265, 6312-6317, 1990  
A:Title: In vivo and in vitro noncovalent association of excised alpha1(I) amino-termi  
A:Reference number: A35233; MUID:90202908; PMID:2318855  
A:Accession: A35233  
A:Molecule type: protein  
A:Residues: 33-52 <WIR>  
A:Note: this propeptide fragment remained non-covalently bound to a defective, uncleaved  
R:Well, D.; d'Alessio, M.; Ramirez, F.; de Wet, W.G.; Chan, D.; Bateman, J.F.

EMBO J. 8, 1705-1710, 1989  
A>Title: A base substitution in the exon of a collagen gene causes alternative splicing  
A:Reference number: S09400; MUID:8935664; PMID:2767050  
A:Accession: S09400  
A:Molecule type: mRNA  
A:Residues: 156-183 <WEI>  
R:Click, E.M.; Bornstein, P.  
Biochemistry 192, 4699-4706, 1970  
A>Title: Isolation and characterization of the cyanogen bromide peptides from the alpha  
A:Reference number: A90567; MUID:71038625; PMID:5529814  
A:Contents: CNBr0-1, CNBr2, CNBr4, CNBr5  
A:Accession: B90567  
A:Molecule type: Protein  
A:Residues: 162-198, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233, 'Z'  
A:Experimental source: skin  
A>Note: evidence for 170-allysine  
R:Beyle, B.; Nodbold, H.; Diebold, J.; Lehmann, H.; Bodo, M.; Deutzmann, R.; Mueller, F.  
Eur. J. Biochem. 192, 153-159, 1990  
A>Title: A critical crosslink region in human bone-derived collagen type I. Specific cle  
A:Reference number: S11372; MUID:90382436; PMID:2169412  
A:Accession: S11372  
A:Molecule type: Protein  
A:Residues: 175-187, 274-287, 'P', 289 <BAE>  
A>Note: sequence of collagen alpha 1(I) isolated from bone after pepsin digestion  
R:Deak, S.B.; Scholz, P.M.; Amenta, P.S.; Constantinou, C.D.; Levi-Mizel, S.A.; Gonzalez  
J. Biol. Chem. 266, 21827-21837, 1991  
A>Title: The substitution of arginine for glycine 85 of the alpha 1(I) procollagen chain  
A:Reference number: I55342; MUID:92042092; PMID:1718984  
A:Accession: I55342  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 258-268, 1347-1357 <DEA>  
A:Cross-references: GB:S67495; NID:G239007; PIDN:AA20350.1; PID:G239008  
A>Note: sequences from the 5' and 3' ends only are shown; mutant sequence 263-Arg report  
R:Morgan, P.H.; Jacobs, H.G.; Segrest, J.P.; Cunningham, L.W.  
J. Biol. Chem. 245, 5042-5048, 1970  
A>Title: Comparative study of glycoproteins derived from selected vertebrate collagens.  
A:Reference number: A92069; MUID:71001508; PMID:4319110  
A:Accession: A92069  
A:Molecule type: Protein  
A:Residues: 263-268 <MOR>  
A:Experimental source: skin  
A>Note: attachment of 2-O-alpha-D-glucosyl-O-beta-D-galactose to 5-hydroxylysine  
R:Labhard, M.E.; Hollister, D.W.  
Matrix 10, 124-130, 1990  
A>Title: Segmental amplification of the entire helical and telopeptide regions of the CD  
A:Reference number: S15989; MUID:90326017; PMID:2374517  
A:Accession: S15989  
A:Molecule type: mRNA  
A:Residues: 281-302, 402-420, 823-843, 925-944, 1026-1045, 1143-1162 <LAB>  
R:Wirtz, M.K.; Rao, V.H.; Glanville, R.W.; Labhard, M.E.; Pretorius, P.J.; de Vries, W.N.  
Connect. Tissue Res. 29, 1-11, 1993  
A>Title: A cysteine for glycine substitution at position 175 in an alpha 1 (I) chain of  
A:Reference number: I52905; MUID:93339042; PMID:8339541  
A:Accession: I52905  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 342-352, 'C', 354-359 <W12>  
A:Cross-references: GB:S64717; NID:G408195; PIDN:AA27677.1; PID:G408196  
A>Note: mutant sequence from patient with osteogenesis imperfecta  
R:Bernard, M.P.; Chu, M.L.; Myers, J.C.; Ramirez, F.; Eikenberry, E.F.; Prockop, D.J.  
Biochemistry 22, 5213-5223, 1983  
A>Title: Nucleotide sequences of complementary deoxyribonucleic acids for the proalpha1  
A:Reference number: A90476; MUID:84080385; PMID:6689127  
A:Accession: A90476  
A:Molecule type: mRNA  
A:Residues: 425-1250, 'X', 1252-1328, 'S', 1330-1390, 'X', 1392-1464 <BR>  
A:Cross-references: GB:K01228; NID:G180391; PIDN:AA51995.1; PID:G180392  
A>Note: sequence partially completed for missing nucleotides by A25439  
R:Chu, M.L.; Garzullo, V.; Williams, C.D.; Ramirez, F.  
J. Biol. Chem. 260, 691-694, 1985  
A>Title: Multixon deletion in an osteogenesis imperfecta variant with increased type I

A:Reference number: A22161; MUID:85104934; PMID:2981843  
A:Accession: A22161  
A:Molecule type: DNA  
A:Residues: 472-594, 'R', 596-607 <CH3>  
A:Cross-references: GB:K03178; GB:K03179; NID:G179612; NID:G179613; PIDN:AA51847.1; PI  
A>Note: the authors translated the codon CGT for residue 595 as Pro  
R:Wallis, G.A.; Starman, B.J.; Zimm, A.B.; Byers, P.H.  
Am. J. Hum. Genet. 46, 1034-1040, 1990  
A>Title: Variable expression of osteogenesis imperfecta in a nuclear family is explained  
A:Reference number: A35336; MUID:90252792; PMID:2339700  
A:Accession: A35336  
A:Molecule type: DNA  
A:Residues: 710-720, 'E', 722-737, 'E', 739-745 <MAL>  
A>Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu  
R:Forlino, A.; Zolazzi, F.; Valli, M.; Pignatelli, P.F.; Cetta, G.; Brunelli, P.C.; Motte  
Hum. Mol. Genet. 3, 2201-2206, 1994  
A>Title: Severe (type III) osteogenesis imperfecta due to glycine substitutions in the  
A:Reference number: I54365; MUID:95187161; PMID:7881420  
A:Accession: I54365  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 746-766, 'S', 768-781 <FOR>  
A:Cross-references: GB:I47667; NID:G1009093; PIDN:AA59576.1; PID:G1009094  
R:Chesler, S.D.; Wallis, G.A.; Byers, P.H.  
J. Biol. Chem. 268, 18218-18225, 1993  
A>Title: Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of ty  
A:Reference number: A47426; MUID:93352646; PMID:8349697  
A:Accession: A47426  
A:Molecule type: mRNA  
A:Residues: 1179-1276, 'H', 1278-1336, 1339-1387, 'R', 1389-1464 <CH>  
A:Cross-references: GB:S64596; NID:G407589; PIDN:AA27856.1; PID:G407590  
A>Note: sequence extracted from NCBI backbone (NCBIN:136444, NCBIPI:136445)  
A:Accession: B47426  
A:Molecule type: mRNA  
A:Residues: 1179-1464 <CH4>  
A:Experimental source: normal dermal fibroblast culture  
A:Accession: C47426  
A:Molecule type: mRNA  
A:Residues: 1179-1276, 'H', 1278-1464 <CH5>  
A:Experimental source: fetal cell 86-237  
A:Accession: D47426  
A:Molecule type: mRNA  
A:Residues: 1179-1336, 1339-1464 <CH6>  
A:Experimental source: fetal cell 86-146  
A:Accession: E47426  
A:Molecule type: mRNA  
A:Residues: 1179-1387, 'R', 1389-1464 <CH7>  
A:Experimental source: fetal cell 88-251  
R:Chou, D.H.; Apone, S.; Byre, D.R.; Starman, B.J.; Andreassen, P.; Charbonneau, H.; Nic  
J. Biol. Chem. 263, 14605-14607, 1988  
A>Title: Substitution of Cysteine for Glycine within the Carboxyl-terminal Telopeptide  
A:Reference number: I55269; MUID:89008319; PMID:3170557  
A:Accession: I55269  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1187-1194, 'C', 1196-1220 <COH>  
A:Cross-references: GB:M23213; NID:G340842; PIDN:AA59363.1; PID:G499622  
A>Note: mutant sequence from a patient with mild osteogenesis imperfecta  
R:Maekelae, J.K.; Raasinen, M.; Vitre, A.; Vuorio, E.  
Nucleic Acids Res. 16, 349, 1988  
A>Title: Human pro-alpha-1(I) collagen: cDNA sequence for the C-propeptide domain.

Query Match 9.8%; Score 127.5; DB 1; Length 1464;  
Best Local Similarity 37.4%; Pred. No. 0.0045;  
Matches 34; Conservative 2; Mismatches 22; Indels 33; Gaps 4;  
QY 3 PGGPAAAS--PQRLRGILLILLIQLPAPSSASRTPGKQKAGLRQRGVVDVYNGMCLQCPA 60  
DB 357 PGGPAAAS--PQRLRGILLILLIQLPAPSSASRTPGKQKAGLRQRGVVDVYNGMCLQCPA 385  
QY 61 GVGGRDGPAGANGIPOTPIGPDGFKBKG 91



A:Residues: 1427-1444 <BRR>  
A:Note: sequence extracted from NCBI backbone (NCBIRP.133363); sequence incorrectly ident  
R:Morrisson, K.E., Matiyama, M.; Yang-Feng, T.L.; Redders, S.T.  
Am. J. Hum. Genet. 49, 545-554, 1991  
A:Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of  
A:Reference number: A39786; MUID:91353570; PMID:1882840  
A:Accession: A39786  
A:Molecule type: mRNA  
A:Residues: 1453-1593, 'A', 1595-1670 <NOR>  
A:Cross-references: GB:S55790; NID:g234418; PIDN:AAAB19637.1; PID:g234419  
A:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope  
C:Genetics:  
A:Gene: COL4A3  
A:Cross-references: GDB:128351; OMIM:120070  
A:Map position: 2q36-2q37  
A:Intons: 1385/1, 1418/1, 1488/1, 1547/2, 1585/3, 1643/2 #status incomplete  
A:Mutons: 1385/1, 1418/1, 1488/1, 1547/2, 1585/3, 1643/2 #status incomplete  
A:Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands with  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 3  
among times: amino-terminal domains (with disulfide and desmosine cross-links), dimeric  
er associations in the interrupted helical domain (with disulfide and desmosine cross-  
C:Function:  
A:Description: minor structural component of extracellular basement membrane in kidney  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracel  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MAT>  
F:29-42/Domain: amino-terminal nonhelical, NH1 <NHI>  
F:43-1438/Region: interrupted helical  
F:791-793/Region: cell attachment (R-G-D) motif  
F:966-998/Region: cell attachment (R-G-D) motif  
F:1154-1156/Region: cell attachment (R-G-D) motif  
F:1306-1308/Region: cell attachment (R-G-D) motif  
F:1345-1347/Region: cell attachment (R-G-D) motif  
F:1432-1434/Region: cell attachment (R-G-D) motif  
F:1439-1670/Domain: carboxyl-terminal nonhelical, NC1 <NC1>  
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CTR>  
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CTR>  
F:31,33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status predi  
F:253/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:1460-1548,1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted  
F:1505-1511,1616-1622/Disulfide bonds: #status predicted  
F:1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted  
Query Match 9.6%; Score 125; DB 1; Length 1670;  
Best Local Similarity 36.2%; Pred. No. 0.0086;  
Matches 34; Conservative 6; Mismatches 36; Indels 18; Gaps 3;  
QY 5 GPAASPRRLRLGLLLLLQLPAPSSASEIRPKQKQALQRREVVLDYNGMCLGGPAGVPG 64  
DB 467 GPPGEP---GLLCTCQCPYIPGPGPLGLGLGVKQIPR-----GGAAGLKG 511  
QY 65 RDGSPGANGIPGTGPIGRD---GFKGEKCELR 95  
DB 512 SPGSPGNTGIPGPFPGAGAGDPLKGEKKEITLQ 545  
RESULT 9  
CGHUTL  
collagen alpha 1(III) chain precursor - human  
N:Alternate names: procollagen alpha 1(III) chain  
C:Species: Homo sapiens (man)  
C:Date: 24-Apr-1984 #sequence revision 01-Sep-1995 #text change 09-Jul-2004  
C:Accession: S05272; S04642; PE0011; S01726; S04887; A90399; A94562; I51868; S59511; A90  
R:Procop, D.J.  
submitted to the EMBL Data Library, February 1989  
A:Reference number: S05272  
A:Accession: S05272  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1240, 'V', 1242-1466 <PRC>  
A:Cross-references: UNIPROT:P02461; EMBL:X14420; NID:g30057; PIDN:CAA32583.1; PID:g30058

R:Ala-Kokko, L.; Kontunaaari, S.; Baldwin, C.T.; Kuivaniemi, H.; Prockop, D.J.  
Biochem. J. 260, 509-516, 1989  
A:Title: Structure of cDNA clones coding for the entire prepro-alpha1(III) chain of hum  
erences.  
A:Reference number: S04642; MUID:89350838; PMID:2764886  
A:Accession: S04642  
A:Molecule type: mRNA  
A:Residues: 1-1196 <ALA>  
A:Cross-references: EMBL:X14420; NID:g30057; PIDN:CAA32583.1; PID:g30058  
A:Note: the complete sequence is not shown  
R:Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.  
Gene 78, 255-265, 1989  
A:Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene  
A:Reference number: PE0011; MUID:89378752; PMID:2777083  
A:Accession: PE0011  
A:Molecule type: DNA  
A:Residues: 1-176 <BEN>  
A:Cross-references: GB:M26939; NID:g180813; PIDN:AAAS2040.1; PID:g180814  
R:Roman, P.D.; Ricci, G.A.; de Crombrughe, B.  
Nucleic Acids Res. 16, 7201, 1988  
A:Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human pr  
A:Reference number: S01726; MUID:88303360; PMID:3405773  
A:Accession: S01726  
A:Molecule type: mRNA  
A:Residues: 1-170 <ROM>  
A:Cross-references: EMBL:X07240; NID:g30060; PIDN:CAA30229.1; PID:g30061  
A:Note: the authors translated the codon CAG for residue 154 as His  
R:Janeczko, R.A.; Ramirez, F.  
Nucleic Acids Res. 17, 6742, 1989  
A:Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen  
A:Reference number: S04887; MUID:89386015; PMID:2780304  
A:Accession: S04887  
A:Molecule type: mRNA  
A:Residues: 149-163, 'G', 164-240, 'D', 242-471, 'D', 473-487, 'L', 489, 'S', 491-613, 'Y', 615-634  
A:Cross-references: EMBL:X15332; NID:g29545; PIDN:CAA33387.1; PID:g930045  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 16, 1158-1164, 1977  
A:Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptid  
A:Reference number: A90399; MUID:77134724; PMID:557355  
A:Accession: A90399  
A:Molecule type: protein  
A:Residues: 'V', 169-225, 229-232, 'P', 234-292, 'D', 294-398 <SEY1>  
A:Experimental source: liver  
A:Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galact  
R:Sever, J.M.  
submitted to the Atlas, December 1977  
A:Reference number: A94562  
A:Accession: A94562  
A:Molecule type: protein  
A:Residues: 'V', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>  
A:Experimental source: liver  
A:Note: author submitted corrections to A90399  
R:MLewicz, D.M.; Wiltz, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P.H.  
Am. J. Hum. Genet. 53, 62-70, 1993  
A:Title: Parental somatic and germ-line mosaicism for a multiexon deletion with unusual  
leping.  
A:Reference number: I51868; MUID:93304430; PMID:8317500  
A:Accession: I51868  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 186-194 <MLT>  
A:Cross-references: GB:S62925; NID:g386425; PIDN:AAD13937.1; PID:g4261637  
R:Chido, A.A.; Silence, D.O.; Cole, W.G.; Bateman, J.F.  
Biochem. J. 311, 939-943, 1995  
A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3  
A:Reference number: S59511; MUID:96067614; PMID:7487954  
A:Accession: S59511  
A:Molecule type: mRNA  
A:Residues: 302-423 <CHI>  
A:Cross-references: GB:S79877; NID:g1195576; PIDN:AAAB35615.1; PID:g1195577  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 17, 3404-3411, 1978

A>Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBr p  
A:Reference number: A90414; PMID:79000343; PMID:667591  
A:Accession: A90414  
A:Molecule type: protein  
A:Residues: 399-675, 'N', 677-727 <SEY3>  
A:Experimental source: liver  
R:Lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.  
J. Biol. Chem. 266, 5256-5259, 1991  
A>Title: G to T transversion at position +5 of a splice donor site causes skipping of th  
A:Reference number: I55349; PMID:91161621; PMID:1672129  
A:Accession: I55349  
A:Molecule type: DNA  
A:Status: translated from GB/EMBL/DBJ  
A:Residues: 537-605 <LEB>  
A:Cross-references: GB:M59312; NID:G180815; PIDN:AAA52041.1; PID:G180816  
R:Sever, J.M.; Mainardi, C.; Kang, A.H.  
Biochemistry 19, 1583-1589, 1980  
A>Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-C85 from ty  
A:Reference number: A90438; PMID:80198282; PMID:6246925  
A:Accession: A90438  
A:Molecule type: protein  
A:Residues: 728-895, 'A', 897-964 <SEY4>  
A:Experimental source: liver  
R:Cole, W.G.; Chiodo, A.A.; Lemande, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Char  
J. Biol. Chem. 265, 17070-17077, 1990  
A>Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping an  
A:Reference number: A38303; PMID:91009133; PMID:2145268  
A:Accession: A38303  
A:Molecule type: mRNA  
A:Residues: 861-1015 <COL>  
A:Cross-references: GB:J05617; GB:M55603; GB:M59227; NID:G180878; PIDN:AA59383.1; PID:G  
A:Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syn  
R:Manooch, B.S.; Dalgleish, R.  
Nucleic Acids Res. 16, 2337, 1988  
A>Title: Human pro alpha1(III) collagen: cDNA sequence for the 3' end.  
A:Reference number: S02119; PMID:88189827; PMID:3357782  
A:Accession: S02119  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 950-1018, 'V', 1020-1183, 'S', 1185-1466 <MAN>  
A:Cross-references: EMBL:X06700; NID:G30053; PIDN:CAA29886.1; PID:G30054  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 20, 2621-2627, 1981  
A>Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-C89 from ty  
A:Reference number: A90446; PMID:81208139; PMID:7016180  
A:Accession: A90446  
A:Molecule type: protein  
A:Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 1157-  
A:Experimental source: liver  
R:Gold, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.; Mye  
Nucleic Acids Res. 12, 9383-9394, 1984  
A>Title: Molecular cloning and carboxyl-propeptide analysis of human type III procollagen  
A:Reference number: A93551; PMID:85087944; PMID:6096827  
A:Accession: A93551  
A:Molecule type: mRNA  
A:Residues: 1065-1155, 'P', 1157-1466 <LOT>  
A:Cross-references: EMBL:X01655; EMBL:X01742; NID:G29584; PIDN:CAA25821.1  
R:Wiskuln, M.; Dalgleish, R.; Kluge-Beckerman, B.; Remard, S.I.; Tolstoshev, P.; Brant  
Biochemistry 25, 1408-1413, 1986  
A>Title: Human type III collagen gene expression is coordinately modulated with the type  
A:Reference number: I52393; PMID:86187804; PMID:3754462  
A:Accession: I52393  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1161-1200 <MIS>  
A:Cross-references: GB:M13146; NID:G180415; PIDN:AAA52003.1; PID:G180416  
R:Manuel, B.S.; Camilizaro, L.A.; Sever, J.M.; Myers, J.C.  
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985  
A>Title: Human alpha 1(III) and alpha 2(IV) procollagen genes are located on the long arm  
A:Reference number: I59025; PMID:85216505; PMID:3658826  
A:Accession: I79359  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

A:Residues: 1165-1196 <EMA>  
A:Cross-references: GB:M11134; NID:G180417; PIDN:AAA52004.1; PID:G180418  
R:Chu, M.L.; Weil, D.; de Wet, W.; Bernard, M.; Sipola, M.; Ramirez, F.  
J. Biol. Chem. 260, 4357-4363, 1985  
A>Title: Isolation of cDNA and genomic clones encoding human pro-alpha1(III) collagen. P  
A:Reference number: A92516; PMID:85157600; PMID:2579949  
A:Accession: A92516  
A:Molecule type: DNA  
A:Residues: 1176-1240, 'V', 1242-1356, 'P', 1358-1466 <CHU>  
A:Cross-references: GB:M10613; GB:M10793; GB:M10794; GB:M10795; GB:M10796; GB:M10797; GB  
A:Experimental source: liver  
A:Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given f  
action  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit ('  
3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently O  
C:Genetics:  
A:Gene: GDB:COL3A1  
A:Cross-references: GDB:118729; OMIM:120180  
A:Map position: 2q31-2q31  
A:Introns: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/3; 1275/1; 1337/3; 1418/3  
A:Note: the list of introns is incomplete; defects in this gene can result in Ehlers-Dan  
C:Complex: type III collagen is a homotrimer of monomers initially linked by disulfide b  
er of their length, is formed with desmosine cross-links made from lysine and allysine r  
C:Function:  
A:Description: structural component of extracellular fibrous polymer that maintains inte  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C:Keywords: coiled coil, Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hyd  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-153/Domain: amino-terminal propeptide #status predicted <PRO>  
F:151-91/Domain: von Willebrand factor type C repeat homology <WVC>  
F:154-1221/Product: collagen alpha 1(III) chain #status predicted <MAT>  
F:154-167/Region: amino-terminal nonhelical telopeptide  
F:168-1196/Region: helical  
F:1091-1093/Region: cell attachment (R-G-D) motif  
F:1197-1221/Region: carboxyl-terminal nonhelical telopeptide  
F:1222-1466/Domain: carboxyl-terminal propeptide #status predicted <CPR>  
F:1238-1466/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:153-154/Cleavage site: Pro-Gln (procollagen N-endopeptidase) #status predicted  
F:154/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicte  
F:161-1212/Modified site: allysine (Lys) #status predicted  
F:163-284, 860, 977, 1106/Modified site: 5-hydroxylysine (Lys) #status experimental  
F:263/Binding site: carboxylate (Lys) (covalent) #status experimental  
F:584, 1094/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental  
F:948-949/Cleavage site: Gly-Ile (collagenase) #status experimental  
F:1106/Binding site: carboxylate (Lys) (covalent) #status predicted  
  
Query Match 9.5%; Score 124; DB 1; Length 1466;  
Best Local Similarity 37.0%; Pred. No. 0.0092;  
Matches 40; Conservative 8; Mismatches 46; Indels 14; Gaps 5;  
  
QY 22 LQLPAPSSASRPPKQKAKAQRREVVDLY--NGWCLGPAVPGDPSPGANGIPRPG 79  
DB 72 LDCPNP---LTPFGCAVCPDPPTAPTRPPNGGPGPKDPPPGIPGNBPGIIPG 127  
  
QY 80 IPRGPGFKGKGECKRESF---EBSWTPNYKCCSSSNLYGIDLGKIA 124  
DB 128 QPGSGSGSGCPPTIC--EBCPTGPGQNYSPQYSDYDKS---GVAAGGLA 170  
  
RESULT 10  
S16366  
collagen alpha 2(IV) chain precursor - pig roundworm  
C:Species: Aegaris suum (pig roundworm)  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: S16366  
R:Petit, J.; Kingston, I.B.  
J. Biol. Chem. 266, 16149-16156, 1991  
A>Title: The complete primary structure of a nematode alpha-2(IV) collagen and the parti  
A:Reference number: S16366; PMID:91340768; PMID:1771907  
A:Accession: S16366  
A:Molecule type: mRNA  
A:Residues: 1-1763 <JBI>





R.Zhou, J.; Hertz, J.M.; Leinonen, A.; Trygsvaason, K.  
J. Biol. Chem. 267, 12475-12481, 1992  
A:Title: Complete amino acid sequence of the human alpha-5(IV) collagen chain and identification of the Alport syndrome patient.  
A:Reference number: S22917; PMID:92316923; PMID:1352287  
A:Accession: S22917  
A:Molecule type: mRNA  
A:Residues: 1-967 <ZHO>  
A:Cross-references: UNIPROT:P29400; GB:M90464; NID:g180826; PIND:AAA52046.1; PID:g553234  
R.Zhou, J.; Leinonen, A.; Trygsvaason, K.  
J. Biol. Chem. 269, 6608-6614, 1994  
A:Title: Structure of the human type IV collagen COL4A5 gene.  
A:Reference number: A54365; PMID:94165049; PMID:8120014  
A:Accession: A54365  
A:Molecule type: DNA  
A:Residues: 1-922 <ZHZ>  
A:Cross-references: GB:U04470; NID:9463378; GB:U04520; NID:9463428; PIND:AA027816.1; PID:R.Zhou, J.; Mochizuki, T.; Smeets, H.; Antignac, C.; Laurila, P.; de Paep, A.; Trygsvaason, K.  
Science 261, 1167-1169, 1993  
A:Title: Deletion of the paired alpha5(IV) and alpha6(IV) collagen genes in inherited sm  
A:Reference number: A57079; PMID:93361972; PMID:8356449  
A:Accession: A57079  
A:Molecule type: DNA  
A:Residues: 1-27 <ZHD>  
A:Cross-references: PDB:237153; NID:9587203; PIND:CAA8512.1; PID:g587204  
R.Philanjanti, T.; Pohjola, E.R.; Myers, J.C.  
J. Biol. Chem. 265, 13758-13766, 1990  
A:Title: Complete primary structure of the triple-helical region and the carboxyl-termin  
A:Reference number: A37122; PMID:90337990; PMID:2380186  
A:Accession: A37122  
A:Molecule type: mRNA  
A:Residues: 84-439, 'GS', '442-624', 'LALG', '629-666', 'FR', '669-887', 'R', '889-1264', '1271-1691' <PIH>  
A:Cross-references: GB:U05558; EMBL:M58526; NID:g3134209  
A:Note: Submitted to the EMBL Data Library, February 1991  
A:Note: the authors translated the codon GCC for residue 115 as Val  
R.Rendert, A.; Serri, M.; Myers, J.C.; Philanjanti, T.; Massella, L.; Rizzoni, G.; De Ma  
Hum. Mol. Genet. 1, 127-129, 1992  
A:Title: De novo mutation in the COL4A5 gene converting glycine 325 to glutamic acid in  
A:Reference number: I54317; PMID:93244772; PMID:1363780  
A:Accession: I54317  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 313-324, 'E', '326-330' <REN>  
A:Cross-references: GB:55934; NID:9299946; PIND:AA013909.1; PID:g4261609  
R.Hoslika, S.L.; Eddy, R.L.; Byers, M.G.; Hoehly, M.; Shows, T.B.; Trygsvaason, K.  
Proc. Natl. Acad. Sci. U.S.A. 87, 1606-1610, 1990  
A:Title: Identification of a distinct type IV collagen alpha chain with restricted kidne  
A:Reference number: A34850; PMID:90160375; PMID:1689491  
A:Accession: A34850  
A:Molecule type: mRNA  
A:Residues: 914-1264, '1271-1691' <HOS>  
A:Cross-references: EMBL:M1115; NID:g180824; PIND:AAA52045.1; PID:g180825  
R.Zhou, J.; Hoslika, S.L.; Chow, L.T.; Trygsvaason, K.  
Genomics 9, 1-9, 1991  
A:Title: Characterization of the 3' half of the human type IV collagen alpha-5 gene that  
A:Reference number: A37969; PMID:91169491; PMID:2004755  
A:Accession: A37969  
A:Molecule type: DNA  
A:Residues: 924-1264, '1271-1691' <ZHZ>  
A:Cross-references: EMBL:M63455; EMBL:M63456; EMBL:M63457; EMBL:M63458; EMBL:M63459; EMB  
8; EMBL:M63470; EMBL:M63471; EMBL:M63472; EMBL:M63473; NID:g177922; PIND:AAA51558.1; PID  
R.Guo, C.; Van Damme, B.; Van Damme-Lombaerts, R.; Van den Berghe, H.; Cassiman, J.J.; M  
Kidney Int. 44, 1316-1321, 1993  
A:Title: Differential splicing of COL4A5 mRNA in kidney and white blood cells: a complex  
A:Reference number: I56971; PMID:9413540; PMID:8301933  
A:Accession: I56971  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1258-1276 <GUO>  
A:Cross-references: GB:569168; NID:9545095; PIND:AA060612.1; PID:g545096  
A:Note: Kidney splice form  
A:Accession: I76598  
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1284-1291, 'TRIGLYACTV' <GUO2>  
A:Cross-references: GB:569169; NID:9545097; PIND:AA060613.1; PID:g545098  
A:Note: frameshift mutation in patient with Alport syndrome  
R.Myers, J.C.; Jones, T.A.; Pohjola, E.R.; Kadri, A.S.; Goddard, A.D.; Sheer, D.; So  
Am. J. Hum. Genet. 46, 1024-1033, 1990  
A:Title: Molecular cloning of alpha5(IV) collagen and assignment of the gene to the regi  
A:Reference number: A35335; PMID:90252791; PMID:2339699  
A:Accession: A35335  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1448-1477 <MYE>  
R.Nakazato, H.; Hattori, S.; Ushijima, T.; Matsura, T.; Koitabashi, Y.; Takada, T.; Yos  
Kidney Int. 46, 1307-1314, 1994  
A:Title: Mutations in the COL4A5 gene in Alport syndrome: a possible mutation in primor  
A:Reference number: I56975; PMID:95156893; PMID:7853788  
A:Accession: I56975  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1595-1602 <NKK>  
A:Cross-references: GB:S75903; NID:9913882; PIND:AA833374.1; PID:g913883  
A:Note: premature termination mutation from a patient with Alport syndrome; one other mu  
R.Lemink, H.H.; Schroeder, C.H.; Brunner, H.G.; Nelen, M.R.; Zhou, J.; Trygsvaason, K.;  
Genomics 17, 485-489, 1993  
A:Title: Identification of four novel mutations in the COL4A5 gene of patients with Alpo  
A:Reference number: I54188; PMID:94010948; PMID:8406498  
A:Accession: I54188  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1604-1607, 'VHDVAKC' <LEM>  
A:Cross-references: GB:S65767; NID:9425563; PIND:AA013967.1; PID:g4261667  
A:Note: frameshift mutation from a patient with Alport syndrome; five other mutations ar  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit ('  
ed and subsequently O-glycosylated).  
C:Genetics:  
A:Gene: GDB:COL4A5; AFS  
A:Cross-references: GDB:120596; OMIM:303630  
A:Map position: Xq22-Xq22  
A:Intons: 27/3; 47/3; 77/3; 92/3; 107/3; 128/3; 146/3; 155/3; 182/3; 203/3; 215/3; 229/  
/3; 799/1; 837/1; 893/1; 923/1; 973/1; 1006/1; 1036/1; 1082/3; 1125/1; 1152/1; 1185/1; 1  
A:Note: the alpha 5(IV) and alpha 6(IV) chain genes are encoded on opposite strands with  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 5(  
mory trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a  
er associations in the interrupted helical domain (with disulfide and desmosine cross-li  
C:Function:  
A:Description: minor structural component of extracellular basement membrane  
A:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: Alport syndrome; basement membrane; coiled coil; extracellular matrix; glyco  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-1691/Product: collagen alpha 5(IV) chain, renal splice form #status predicted <MAT1  
F:27-1264,1271-1691/Product: collagen alpha 5(IV) chain, leukocyte splice form #status p  
F:77-41/Domain: amino-terminal nonhelical, NC2 #status predicted <NC2>  
F:1462-1462/Region: interrupted helical  
F:1463-1691/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>  
F:1473-1573/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F:1583-1697/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F:29-32,38-40,124-451,481-484/Disulfide bonds: interchain #status predicted  
F:125/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:1482-1570,1515-1573/Disulfide bonds: (or 1482-1573, 1515-1570) #status predicted  
F:1522-1533,1638-1644/Disulfide bonds: #status predicted  
F:1592-1684,1626-1687/Disulfide bonds: (or 1592-1687, 1626-1684) #status predicted

Query Match 9.4%; Score 123; DB 1; Length 1691;  
Best local Similarity 36.5%; Pred. No. 0.013; Indels 16; Gaps 4;  
Matches 35; Conservative 7; Mismatches 38;

QY 3 PGAPAA--SPQRLRGILLILLQLPAPSSASRIP-----KGKQKALQRREVVLDLYNQC 55  
DB 1256 PGGPGRRPRPRPGG-----LPGPGRRPGLPENGKIGK-EKKNPPQGLPGLGLK 1306

QY 56 LGGPAGVGRDSSPGANGIPGTPGIPGRDGFKEKRG 91  
DB 1307 DGGPPGLQGNPGRPGANGKGGPGLPGVDFPGMKG 1342

## RESULT 15

T24482  
hypothetical protein T05A1.2 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T24482  
R,Lloyd, C.  
submitted to the EMBL Data Library, December 1995  
A/Reference number: Z19897  
A/Accession: T24482  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-300 <WIL>  
A/Cross-references: UNIPROT:Q22183; EMBL:Z68219; PIDN:CAA92476.1; GSPDB:GN00022; CESP:T05A1  
A/Experimental source: clone T05A1  
C/Genetics:  
A/Gene: CESP:T05A1.2  
A/Map position: 4  
A/Introns: 7/3

Query Match 9.4%; Score 122.5; DB 2; Length 300;  
Best Local Similarity 31.5%; Pred. No. 0.0021;  
Matches 35; Conservative 11; Mismatches 28; Indels 37; Gaps 5;  
QY 3 PGGPAASPQR--LRGLLLILLQLPAPSSASEIPKGRKQKQLRQREVVDLYNGMCLQGP- 59  
Db 189 PHGPNGHPGKPKGSQG-----PGPPGHSDPEKPKQPGQPGRAGP-----RGPR 231  
QY 60 --AGVPGRGSSPGA-----NGIPGTGIRGRDGFKGRKGRFC 93  
Db 232 GVAGIRKGDGAPSPGQPGPRGGPGEPPGDGAPQGPGLPSDGLPGRKGLC 282

Search completed: August 4, 2005, 04:22:29  
Job time : 20.3765 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:07:24 ; Search time 88.6235 Seconds

(without alignments)  
1404.088 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303  
Sequence: 1 MRQGPASRQRRLGILL.....GDASTGMSVSRILIELPK 243

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	2	096CG8 homo sapien
2	1294	94.3	2	06UW91	06UW91 mus sapien
3	1225	94.0	245	2	08CG08 rattus norv
4	1210	92.9	245	2	09DID6 mus musculu
5	1061	81.4	232	2	081X63 homo sapien
6	928	71.2	231	2	06AX10 brachydanio
7	139	10.7	565	2	08K036 mus musculu
8	133.5	10.2	717	2	09N052
9	131.5	10.1	246	1	CIQC MOUSE
10	131.5	10.1	1472	2	0902A0
11	129	9.9	287	2	08CFR0
12	129	9.9	1752	2	007265
13	128	9.8	289	2	018799
14	127.5	9.8	358	2	06MEY7
15	127.5	9.8	1069	2	06LAN8
16	127.5	9.8	1461	2	076045
17	127.5	9.8	1464	1	CALL_HUMAN
18	127.5	9.8	1464	1	08N473
19	127	9.7	291	2	09NAR3
20	126.5	9.7	540	2	086Y22
21	126.5	9.7	568	2	08CD80
22	126.5	9.7	739	2	070575
23	126.5	9.7	747	2	06NWS7
24	126.5	9.7	751	2	09RIN9
25	126	9.7	327	2	001799
26	126	9.7	754	1	CAS4 CANFA
27	126	9.7	1447	2	06PAU1
28	126	9.7	1447	2	06U135
29	126	9.7	1669	2	09Q2S0
30	125	9.6	551	2	06ZM13
31	125	9.6	886	2	09NUB7

## ALIGNMENTS

RESULT 1	ID	Q96CG8	PRELIMINARY	PRT	243 AA.
AC	Q96CG8				
DT	01-DEC-2001	(Tremblrel. 19, Created)			
DT	01-DEC-2001	(Tremblrel. 19, Last sequence update)			
DT	25-OCT-2004	(Tremblrel. 28, Last annotation update)			
DE	Collagen triple helix repeat containing 1 (Collagen triple helix repeat-containing protein 1).				
GN	Name=CTHRC1				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Datchenko L., Marsiska K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Utsdin T.B., Toshiyuki S., Carninci P., Prange C., Kana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzyzanski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E., Jones S.J., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				
RC	Strausberg R.J.				
RL	Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Adult smooth muscle;				
RC	Lehner W., Moore D.P., Harmon K.J., Mancini M.L., Lindner V.;				
RL	Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.				
DR	EMBL: BC014245; AA014245.1; -				
DR	EMBL: AY136825; AA015749.1; -				
DR	Gene, HGNC:18631; CTRCL.				
DR	GO: 0005737; C:cytoplasm; IEA.				
DR	GO: 0006817; P:phosphate transport; IEA.				
DR	InterPro: IPR008161; Clg_helix.				

32	125	9.6	888	2	090796	090796 gallus gall
33	125	9.6	1670	1	CAS4_HUMAN	001955 homo sapien
34	125	9.6	1685	1	CAS4_HUMAN	P29400 homo sapien
35	124.5	9.6	289	1	CQT7_HUMAN	Q9BXJ2 homo sapien
36	124.5	9.6	326	2	0677V9	0677V9 lymphocysti
37	124.5	9.6	1336	2	06R241	06R241 homo sapien
38	124.5	9.6	1516	1	CALL_HUMAN	P39060 homo sapien
39	124.5	9.6	1516	2	06R239	06R239 homo sapien
40	124.5	9.6	1684	2	08HYC1	08HYC1 canis fami1
41	124.5	9.6	1691	2	0866Z2	0866Z2 canis fami1
42	124.5	9.6	1751	2	06R240	06R240 homo sapien
43	124	9.5	957	2	096P44	096P44 homo sapien
44	124	9.5	957	2	09H0V3	09H0V3 homo sapien
45	124	9.5	1163	2	08N6U4	08N6U4 homo sapien

DR InterPro: IPR008160; Collagen.

DR Pfam: PF01391; Collagen; 1.

DR ProDom: PD000007; Clg\_helix; 1.

KW Collagen.

SQ SEQUENCE 243 AA; 26224 MW; A11FE81C6687F9 CRC64;

Query Match 100.0%; Score 1303; DB 2; Length 243;

Best Local Similarity 100.0%; Pred. No. 9.1e-107;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPASPORLRGLLLLLQLPAPSSASRIPKQKQAKLRQREVVLDYNGMCLQGPA 60

DB 1 MRPQGPASPORLRGLLLLLQLPAPSSASRIPKQKQAKLRQREVVLDYNGMCLQGPA 60

QY 61 GVPGRDGPANGIPGTGPIGRDGFKEGKECLRESFEESWTPTYKQCSWSLNYGIDL 120

DB 61 GVPGRDGPANGIPGTGPIGRDGFKEGKECLRESFEESWTPTYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDQ 180

DB 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDQ 180

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240

DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 1 MRPQGPASPORLRGLLLLLQLPAPSSASRIPKQKQAKLRQREVVLDYNGMCLQGPA 60

DB 1 MRPQGPASPORLRGLLLLLQLPAPSSASRIPKQKQAKLRQREVVLDYNGMCLQGPA 60

QY 61 GVPGRDGPANGIPGTGPIGRDGFKEGKECLRESFEESWTPTYKQCSWSLNYGIDL 120

DB 61 GVPGRDGPANGIPGTGPIGRDGFKEGKECLRESFEESWTPTYKQCSWSLNYGIDL 120

QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDQ 180

DB 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDQ 180

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240

DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSRITIEE 240

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

DB 241 LPK 243

QY 241 LPK 243

RESULT 4  
 Q9DID6 PRELIMINARY; PRT; 245 AA.  
 AC Q9DID6;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:1110014B07 product:hypothetical collagen triple helix repeat containing protein, full insert sequence.  
 DE Repeat containing protein, full insert sequence.  
 GN Name=Ctrcl;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20493374; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20493374; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P., Kono H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishire T., Harada A., Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matshiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arawaka T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanezaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kaskawa T., Kato H., Kawai T., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai K., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RU Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK03674; BA522930.1; -;  
 DR MGD; MGI:1915838; Ctrcl.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR008161; C1g\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR Pfam; PF01391; Collagen; 1.  
 DR Prodom; PD000007; C1g\_helix; 1.  
 KW Collagen, Hypothetical protein.  
 SQ SEQUENCE 245 AA; 26460 MW; 14951B87D181A0E CRC64;  
 Query Match 92.9%; Score 1210; DB 2; Length 245;  
 Best Local Similarity 93.5%; Pred. No. 1,5e-98;  
 Matches 229; Conservative 4; Mismatches 10; Indels 2; Gaps 1;  
 QY 1 MRPGPAASPORRG--LLLLLLLPAPSSASEIPKQKQALRQREVVDLYNGMCLQG 58  
 DB 1 MRPGPAAPOLLGLFLVLLLLQSAPISSSENKVKQALIRQREVVDLYNGMCLQG 60  
 QY 59 PACVPRDPSPGANGIPGTGPIGRDGFKEGKCELRSEESWTNNYKQCSWSLNYGI 118  
 DB 61 PACVPRDPSPGANGIPGTGPIGRDGFKEGKCELRSEESWTNNYKQCSWSLNYGI 120  
 QY 119 DLGKIAECTFTKRSNSAIRVLPFSGSLRKCNNACCRNYFTTNGACGGLPIEAIYL 178  
 DB 121 DLGKIAECTFTKRSNSAIRVLPFSGSLRKCNNACCRNYFTTNGACGGLPIEAIYL 180  
 QY 179 DQSPENASTINHRISVVEGLCEGIGAGLVDAIVWGTCSDPKPDASTGNMSVRRII 238  
 DB 181 DQSPENASTINHRISVVEGLCEGIGAGLVDAIVWGTCSDPKPDASTGNMSVRRII 240  
 QY 239 EELPK 243  
 DB 241 EELPK 245  
 RESULT 5  
 Q81X63 PRELIMINARY; PRT; 232 AA.  
 AC Q81X63;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NTMC1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sanuki N., Fujiki K., Kanai A., Tanaka Y., Iwata T.;  
 RU Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; A3395488; AAO17919.1; -;  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR008161; C1g\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR Pfam; PF01391; Collagen; 1.  
 DR Prodom; PD000007; C1g\_helix; 1.  
 KW Collagen.  
 SQ SEQUENCE 232 AA; 25163 MW; E9D4BC30304837ED CRC64;  
 Query Match 81.4%; Score 1061; DB 2; Length 232;  
 Best Local Similarity 83.5%; Pred. No. 1,9e-85;  
 Matches 202; Conservative 8; Mismatches 18; Indels 14; Gaps 2;  
 QY 1 MRPGPAASPORRG--LLLLLLLPAPSSASEIPKQKQALRQREVVDLYNGMCLQGA 60  
 DB 1 MRPGPAASPORRG--LLLLLLLPAPSSASEIPKQKQALRQREVVDLYNGMCLQGA 60  
 DB 1 MRPGPAASPORRG--LLLLLLLPAPSSASEIPKQKQALRQREVVDLYNGMCLQGA 60

```
QY 61 GVPGRDGPANGIPGTGGIPGRDGFKEGKECELRSPFESWTPNYKQCSWSLNYGIDL 120
Db 47 GVPGRGSGPANGIPGTGGIPGRDGFKEGKECELRSPFESWTPNYKQCSWSLNYGIDL 106
QY 121 GKIAECTFTKMRNSALRYLFGSGLRLKCRNACCQRMWTFPENGACSGPLPIEAIYLDQ 180
Db 107 GKIAECTFTKMRNSALRYLFGSGLRLKCRNACCQRMWTFPENGACSGPLPIEAIYLDQ 166
QY 181 GSPENNSTINHTSSVEGICGEGIGAGLVDAIWMVGTCSDPYKGDASTGMNSVSRITIEE 240
Db 167 GSPENNSTINHTSSVEGICGEGIGAGLVDAIWMVGTCSDPYKGDASTGMNSVSRITIEE 226
QY 241 LP 242
Db 227 LP 228

RESULT 6
Q6AXL0 PRELIMINARY; PRT; 231 AA.
ID 06AXL0;
AC 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Zgc:101075 protein.
GN Name=zgc:101075;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP STRAIN=Singapore local strain; TISSUE=Embryo;
RC PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RA Director MGC Project;
RA Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC079494; AAH79494.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR InterPro; IPR008167; P:phosphate transport; IEA.
DR DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 1.
KW Collagen.
SQ SEQUENCE 231 AA; 25395 MW; 4F9C698080FE19B CRC64;
```

```
Db 4 KTLQLLICFWISLPFCVYQKAKERIPR-----QDAFTTKYQA-CYQGVVGVGRGN 56
QY 69 PGANGIPGTGGIPGRDGFKEGKECELRSPFESWTPNYKQCSWSLNYGIDLKIAECTF 128
Db 57 PGANGIPGTGGIPGRDGFKEGKECELRSPFESWTPNYKQCSWSLNYGIDLKIAECTF 116
QY 129 TKMRNSALRYLFGSGLRLKCRNACCQRMWTFPENGACSGPLPIEAIYLDQSPENNST 188
Db 117 TKMRNSALRYLFGSGLRLKCRNACCQRMWTFPENGACSGPLPIEAIYLDQSPENNST 176
QY 189 INIHTSSVEGICGEGIGAGLVDAIWMVGTCSDPYKGDASTGMNSVSRITIEE 243
Db 177 INIHTSSVEGICGEGIGAGLVDAIWMVGTCSDPYKGDASTGMNSVSRITIEE 231

RESULT 7
Q8K036 PRELIMINARY; PRT; 565 AA.
ID 08K036;
AC 08K036;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Col13a1 protein.
DE Name=Col13a1;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;
RA Strausberg R.;
RA Submitted (Jul-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC034164; AAH34164.1; -.
DR WGI; WGI:1277201; Col13a1.
DR GO; GO:0005911; C:intercellular junction; IDA.
DR InterPro; IPR008161; C1g_helix.
DR DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 5.
DR ProDom; PD000007; C1g_helix; 1.
KW Collagen.
SQ SEQUENCE 565 AA; 56726 MW; DBD3FF99D670195F CRC64;
```

Query Match 71.2%; Score 928; DB 2; Length 231;  
Best Local Similarity 72.8%; Pred. No. 1e-73;  
Matches 111; Conservative 21; Mismatches 33; Indels 10; Gaps 3;

Query Match 10.7%; Score 139; DB 2; Length 565;  
Best Local Similarity 34.8%; Pred. No. 0.0011;  
Matches 39; Conservative 12; Mismatches 35; Indels 26; Gaps 6;



```

Db      393  PGAAGGCGSPGPG-AKGEPRKGMVD-YNGSINBALQETRTLTALMPGRLPGOTGPRGP 450
QY      72  NGIP-----GTTPGPRDGPKEKGECLRESEESTPNYKCCSWSLNYG 117
Db      451  PGTGPRGGEIGLPGPPGHDGKPRG---KPGDAGWSRTPRKGMASRNREG 499

RESULT 8
Q9N052  PRELIMINARY; PRT; 717 AA.
ID      Q9N052
AC      Q9N052;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Type XIII collagen.
GN      Name=COLXIII1;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Snellman A., Tu H., Vaeisaeen T., Kvist A.P., Huhtala P.,
RA      Pihlajaniemi T.;
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ293624; CAC00688.1; -.
DR      Genew; HGNC:2190; COL13A1.
DR      GO; GO:0005737; Cytoplasm; IEA.
DR      GO; GO:0006817; P:phosphate transport; IEA.
DR      InterPro; IPR008161; C1q helix.
DR      InterPro; IPR008160; Collagen.
DR      Pfam; PF01391; Collagen; 9.
DR      ProDom; PD000007; C1q_helix; 3.
KW      Collagen.
SQ      SEQUENCE 717 AA; 69964 MW; A311B9C7D3B87577 CRC64;

Query Match 10.2%; Score 133.5; DB 2; Length 717;
Best Local Similarity 43.4%; Pred. No. 0.0045;
Matches 33; Conservative 8; Mismatches 16; Indels 19; Gaps 4;

QY      34  PKGKQKALQREVDVLYNG-----MCLQGPAGVPGRDGSGANGIP-----G 76
Db      431  PKG-SKGEPRKGMVD-YNGSINBALQETRTLTALMPGRLPGOTGPRGP 488

QY      77  TPGTPGRDGPKEKGE 92
Db      489  LFGPPGHDGKPRGK 504

RESULT 9
C1QC_MOUSE STANDARD; PRT; 246 AA.
ID      C1QC_MOUSE
AC      Q02105;
DT      01-JUL-1993 (Rel. 26, Created)
DT      01-JUL-1993 (Rel. 26, Last sequence update)
DT      25-JAN-2005 (Rel. 46, Last annotation update)
DE      Complement C1q subcomponent, C chain precursor.
GN      Name=C1qg; Synonyms=C1q;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ICR; TISSUE=Macrophage;
RX      MEDLINE=9301118; PubMed=1396691;
RA      Petry F., Reid K.B.M., Looos M.;
RA      Petry F., Reid K.B.M., Looos M.;
RT      "Isolation, sequence analysis and characterization of cDNA clones
RT      coding for the C chain of mouse C1q. Sequence similarity of complement
RT      subcomponent C1q, collagen type VIII and type X and precerebellin.",
RL      Eur. J. Biochem. 209:129-134(1992).

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=BALB/c; TISSUE=Liver;
RX      MEDLINE=96186528; PubMed=8606057; DOI=10.1007/s002510050077;
RA      Petry F., McClive P.J., Bolto M., Morley B.J., Moran G., Looos M.;
RT      "The mouse C1q genes are clustered on chromosome 4 and show
RT      conservation of gene organization.";
RL      Immunogenetics 43:370-376(1996).

RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Colon;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA      Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahney J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA      Scherch A.A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC      -1- FUNCTION: C1q associates with the proenzymes C1r and C1s to yield
CC      C1, the first component of the serum complement system. The
CC      collagen-like regions of C1q interact with the Ca(2+)-dependent
CC      C1r(2)C1s(2) proenzyme complex, and efficient activation of C1
CC      takes place on interaction of the globular heads of C1q with the
CC      Fc regions of IgG or IgM antibody present in immune complexes.
CC      -1- SUBUNIT: C1 is a calcium-dependent trimolecular complex of C1q, R
CC      and S in the molar ration of 1:2:2. C1q subcomponent is composed
CC      of nine subunits, six of which are disulfide-linked dimers of the
CC      A and B chains, and three of which are disulfide-linked dimers of
CC      the C chain.
CC      -1- SIMILARITY: Contains 1 C1q domain.
CC      -1- SIMILARITY: Contains 1 collagenous domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      EMBL; X66295; CAA46993.1; -.
CC      EMBL; X92960; CAA63535.1; -.
CC      EMBL; BC054443; AAH54443.1; -.
CC      F1R; S29328; S29328.
CC      HSSP; G60994; IC28.
CC      MGD; MGI:88225; C1qg.
CC      InterPro; IPR001073; C1q.
CC      InterPro; IPR008161; C1q helix.
CC      InterPro; IPR008160; Collagen.
CC      InterPro; IPR008983; TNP_like.
CC      Pfam; PF00386; C1q; 1.
CC      DR      Pfam; PF01391; Collagen; 1.
CC      DR      PRINTS; PRO0007; COMPLEMENTC1Q.
CC      DR      ProDom; PD000007; C1q helix; 1.
CC      DR      PROSITE; PS50871; C1Q_1.
CC      KW      Collagen; Complement pathway; Glycoprotein; Hydroxylation; Plasma;
CC      Repeat; Signal.
CC      FT      SIGNAL 1 29
CC      FT      CHAIN 30 246 Complement C1q subcomponent, C chain.
CC      FT      DOMAIN 32 113 Collagen-like.

```

```
FT DOMAIN 116 246 C1q.
FT DISULFID 33 33 Interchain (with other C chain) (By
FT MOD RES 37 37 Hydroxyproline (By similarity).
FT MOD RES 40 40 Hydroxyproline (By similarity).
FT MOD RES 43 43 Hydroxyproline (By similarity).
FT MOD RES 61 61 Hydroxyproline (By similarity).
FT MOD RES 64 64 Hydroxyproline (By similarity).
FT MOD RES 73 73 5-Hydroxylysine (By similarity).
FT MOD RES 76 76 5-Hydroxylysine (By similarity).
FT MOD RES 79 79 Hydroxyproline (By similarity).
FT MOD RES 82 82 Hydroxyproline (By similarity).
FT MOD RES 97 97 Hydroxyproline (By similarity).
FT MOD RES 100 100 Hydroxyproline (By similarity).
FT MOD RES 106 106 Hydroxyproline (By similarity).
FT MOD RES 109 109 Hydroxyproline (By similarity).
SQ SEQUENCE 246 AA; 25966 MW; 2F79EAL274BCB8E0 CRC64;

Query Match 10.1%; Score 131.5; DB 1; Length 246;
Best Local Similarity 38.5%; Pred. No. 0.0019;
Matches 35; Conservative 5; Mismatches 32; Indels 19; Gaps 2;

QY 5 GPAPSPQRRLGILLILLQLPAPSSASEIPKGRKQAQLRQREVVDYNGMCIOGPAGVVG 64
DB 4 GPSCQPCQCGILLFLPLPLRSQS-----AGCYGIPGMGMGADG 47
QY 65 RDGS---PGANGIPGTGIPGRDGFKEGKE 92
DB 48 KDGHDGLQGPKEPGRIPAVPGTQPGKGQGE 78

RESULT 10
Q90ZAO PRELIMINARY; PRT; 1472 AA.
ID Q90ZAO;
DC 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Collagen type XX alpha 1 precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX NCBI_TaxID=9031;
RP SEQUENCE FROM N.A.
RX MEDLINE=21303548; PubMed=11274142; DOI=10.1074/jbc.M009912200;
RA Koch M., Foley J.E., Hahn R., Burgess R.E., Gerecke D.R.,
RA Gordon M.K.;
RT "alpha 1(X) collagen, a new member of the collagen subfamily, fibril-
RT associated collagens with interrupted triple helices.";
RL J. Biol. Chem. 276:23120-23126(2001).
DR EMBL; AF312825; AAK58847.1; -.
DR HSSP; Q96KP7; 1FNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR Pfam; PF01391; Collagen; 4.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR ProDom; PD000007; C1q_helix; 4.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00853; FN3; 6.
DR PROSITE; PS50234; VWA; 1.
KW Collagen; Signal.
FT SIGNAL 1 28 Potential.
SQ SEQUENCE 1472 AA; 156903 MW; 5361611579C56EFD CRC64;

Query Match 10.1%; Score 131.5; DB 2; Length 1472;
```

```
Best Local Similarity 34.5%; Pred. No. 0.016;
Matches 38; Conservative 7; Mismatches 26; Indels 39; Gaps 6;

QY 3 PGAPSPQRRLGILLILLQLP-----APSSASEI---PGKQAQLRQREV 47
DB 1161 PGPGSPGR-RG-----PGQGPGRKPGFPKGPAGPGCGQSSPSQ--- 1207
QY 48 VDLNMGMLQGPAGVP-----GRDSSPANGIPGTGIPGRDGFKEGKE 91
DB 1208 -----GITTIQGVPPGIKEKGTGSPMGQIPGVQAPAGDGLQAGK 1252

RESULT 11
Q9CFR0 PRELIMINARY; PRT; 287 AA.
ID Q9CFR0;
DC 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to C1q-like.
GN Name=BC040774;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Wuzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.T., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RX Submitted (Nov-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040774; AAH40774.1; -.
DR HSSP; Q60994; 1C28.
DR MGD; MGI:3032521; BC040774.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen.
SQ SEQUENCE 287 AA; 29292 MW; 8FF89BC1C7420415 CRC64;

Query Match 9.9%; Score 129; DB 2; Length 287;
Best Local Similarity 35.5%; Pred. No. 0.0039;
Matches 43; Conservative 10; Mismatches 56; Indels 12; Gaps 4;
```

```

QY 27 PSSASEIPKCKOKAQLRQREV-DLYNG---MCIQGPAGVPRDGSFGANGIPGTPTGPG 82
DB 44 PAKAKAPPPSTAL---EVMODLSANPPPIQGGKGPGRKGPGRPGEPGPG 100
QY 83 RQGFKEGCEGLRESFEBSWTNPKQCSGLNYGIDLGIACTFTNMSNGLRVLPFS 142
DB 101 PRPPPEKGDGSGRPGU-----PGLQLTSSAGGVGVSGGTGGGCTEGEVTSLSAAPS 155
QY 143 G 143
DB 156 G 156

RESULT 12
Q07265 PRELIMINARY; PRT; 1752 AA.
ID 007265;
AC 007265;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 3 alpha procollagen.
GN Name=COLP3alpha;
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Echinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93186842; PubMed=8444899;
RA Exposito J.-Y., D'Alessio M., Di Liberto M., Ramirez F.;
RT "Complete primary structure of a sea-urchin type IV collagen and
RT analysis of the 5' end of its gene.";
RL J. Biol. Chem. 268:5249-5254(1993).
DR EMBL; L02917; AAA30039.1; -.
DR HSSP; P02462; 1L11.
DR GO; GO:0005581; C:collagen; IEA.
DR GO; GO:0005737; C:cyclopilasm; IEA.
DR GO; GO:0005201; F:extracellular matrix structural constituent; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001442; Procollagen4_C.
DR Pfam; PF01413; C4; 2.
DR Pfam; PF01391; Collagen; 25.
DR SMART; SM00111; C4; 2.
DR COLLAGEN.
KW COLLAGEN.
SQ SEQUENCE 1752 AA; 170210 MW; 1A55AA21569346D CRC64;

Query Match 9.9%; Score 129; DB 2; Length 1752;
Best Local Similarity 22.9%; Pred. No. 0.032;
Matches 47; Conservative 20; Mismatches 60; Indels 78; Gaps 9;

QY 3 PGGPASPORLRLILLILLQLPAPSSASEIPKCKOKAQLRQREVVDLYNGMCLQSPAGV 62
DB 1453 PGGPSPRDR-----PGPG-----LTGDKGT 1477
QY 63 PGGDSSPGANGIPGTGPIGRDGFKEGCEGLRESF-----ESWT 103
DB 1478 PGVQGNPRVSGVGEGLKGEQFRQNGPGRPGFPGTKGEAGIPGSSSGFTTHSGT 1537
QY 104 PNYKQC-----SW-----SSLNIGIDLGIACTFTNMSNGLRVLPFSGLRL 147
DB 1538 TSPQCPQGTAKMKGYSLLFVQGNRGRGQDLGKPGSC-----LKRFTWPLFC-NINN 1592
QY 148 KCRNACCORWYFTFNGACSGPLPI 172
DB 1593 VCHVASRNDYXWLSLTT---PWPM 1614

RESULT 13
Q18799 PRELIMINARY; PRT; 289 AA.
ID Q18799

```

```

AC Q18799;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C53B4.5.
GN ORFNames=C53B4.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Berks M.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z68215; CAA92453.1; -.
DR PIR; T20177; T20177.
DR WormBase; WBGene0000693; C53B4.5.
DR WormPep; C53B4.5; C603091.
DR GO; GO:0005737; C:cyclopilasm; IEA.
DR GO; GO:0042302; P:structural constituent of cuticle; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002486; Col cuticle_N.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF01484; Col cuticle_N; 1.
DR COLLAGEN; Hypothetical protein.
KW COLLAGEN.
SQ SEQUENCE 289 AA; 28993 MW; 2999A3F9CC0B044 CRC64;

Query Match 9.8%; Score 128; DB 2; Length 289;
Best Local Similarity 35.7%; Pred. No. 0.0048;
Matches 30; Conservative 5; Mismatches 19; Indels 30; Gaps 2;

QY 25 PPSASEIPKCKOKAQLRQREVVDLYNGMCLQSPAGVGRDGS-----P 69
DB 203 PPGSFGEPGRPGQPS-----RGPAGQKQAGCGPGEKANGEPGP 247
QY 70 GANIGTPGPIGRDGFKEGCEGC 93
DB 248 GRDQGRGRCQGRDGRHPEKGV 271

RESULT 14
Q06MEY7 PRELIMINARY; PRT; 358 AA.
ID Q06MEY7;
AC Q06MEY7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedCusNames=pc0138;
OS Parachlamydia sp. (strain UME25) (subsp. Acanthamoeba sp.).
OC Bacteria; Chlamydiae; Parachlamydiae; Parachlamydiaceae; Parachlamydia.
OX NCBI_TaxID=264201;
RN [1]
RP SEQUENCE FROM N.A.
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA Fartmann B., Brandt P., Nyakatura G.J., Droege M., Friehman D.,
RA Rattei T., Mewes H.-W., Wagner M.;
RT "Genome sequence of an amoeba symbiont and its use for reconstructing
RT the evolutionary history of chlamydiae.";
RL Submitted (JAN-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; BX908798; CAF22862.1; -.
DR GO; GO:0005737; C:cyclopilasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.

```

DR InterPro; IPR008160; Collagen.  
DR Pfam; PF01391; Collagen; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 358 AA; 37332 MW; 87419BD361E3D61B CRC64;

Query Match 9.8%; Score 127.5; DB 2; Length 358;  
Best Local Similarity 28.8%; Pred. No. 0.0068;

Matches 46; Conservative 15; Mismatches 72; Indels 27; Gaps 5;

QY 34 PKGKQKQRLRQREVDLYNMGCIQGPAGVPGRDGSPGANGIPGTPGIPGRDGFKEKG-- 91  
DB 136 PPGPQCATGPQGPFGPTGAGLTGPTGLTGATGLTGADGFPADGWSGSKGLT 195  
QY 92 -----ECLRESFEESWTPNY-----KQCSWSLNYGIDLGKIAE-----CTF 128  
DB 196 GATGLTGATGLTGASGPATGANNVFPAYGIETGSVNPVNLSPDVCPIIDGWRPTNTTF 255  
QY 129 TTKRNSNLSRLVLFSGSLRLKCRNACCQRYTFPFGACSG 168  
DB 256 TCKQTGTLYL-VQVRGQFTLVMDATCALW-ATFNGSQVEG 293

## RESULT 15

Q6LAN8 PRELIMINARY; PRT; 1069 AA.  
AC O6LAN8;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Collagen type I alpha 1 (Fragment).  
GN Name=COL1A1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97141927; PubMed=8988177;  
RA Simon M., Pedetour F., O'Brien K., Sirvent N., Kedra D., Guilbaud C.,  
Turc-Carel C., Dunanski J.P.,  
RT "Regulation of the platelet-derived growth factor B-chain gene via  
RT fusion with collagen gene COL1A1 in dermatofibrosarcoma protuberans  
RT and giant-cell fibroblastoma.";  
RL Nat. Genet. 15:95-98(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA O'Brien K.P.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X98705; CAA67261.1; -.  
DR GO; GO:0005737; Cytoplasm; IEA.  
DR GO; GO:0006817; Phosphate transport; IEA.  
DR InterPro; IPR008161; C1g\_helix.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR009041; PMP\_SGCL.  
DR Pfam; PF01391; Collagen; 16.  
DR Pfam; PF00093; WVC; 1.  
DR ProDom; PD000007; C1g\_helix; 3.  
DR SMART; SM00214; WVC; 1.  
DR PROSITE; PS01208; WVC\_1; 1.  
DR PROSITE; PS01084; WVC\_2; 1.  
KW Collagen.  
FT NON TER 1069  
SQ SEQUENCE 1069 AA; 97445 MW; EE279B10572FB980 CRC64;

Query Match 9.8%; Score 127.5; DB 2; Length 1069;  
Best Local Similarity 37.4%; Pred. No. 0.024;

Matches 34; Conservative 2; Mismatches 22; Indels 33; Gaps 4;

QY 3 PGGPAA--PGRKGLLLLLLLLPAPSSASEIPKQKQKQRLRQREVDLYNMGCIQGPA 60  
DB 357 PGGPAA--PGRKGLLLLLLLLPAPSSASEIPKQKQKQRLRQREVDLYNMGCIQGPA 385

QY 61 GVGGRDGPANGIPGTPGIPGRDGFKEKG 91  
DB 386 GNPAGDGPQGAANGANGAPGIAAGPFGARG 416

Search completed: August 4, 2005, 04:21:23  
Job time : 89.6235 secs



```
Db      61 GVPGRGSGPANGIPGTPTGIPGRDGFKEGKEGECLERESFEBSWTPTYKQCSWSLNYGIDL 120
      |||
Qy      121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDO 180
      |||
Db      121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDO 180
      |||
Qy      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
      |||
Db      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
      |||
Qy      241 LPK 243
      |||
Db      241 LPK 243

RESULT 2
US-10-045-992-4
; Sequence 4, Application US/10045992
; Publication No. US20020161211A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert E.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN
; FILE REFERENCE: 053689-5006-01
; CURRENT APPLICATION NUMBER: US/10/045,992
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-045-992-4

Query Match      100.0%; Score 1303; DB 13; Length 243;
Best Local Similarity 100.0%; Pred. No.1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MRQGPASFORLRGLLLLLQLPAPSSASBIPKQKQAKQLRQREVVLDYNGMCLQCPA 60
      |||
Db      1 MRQGPASFORLRGLLLLLQLPAPSSASBIPKQKQAKQLRQREVVLDYNGMCLQCPA 60
      |||
Qy      61 GVPGRDGSFGANGIPGTPTGIPGRDGFKEGKEGECLERESFEBSWTPTYKQCSWSLNYGIDL 120
      |||
Db      61 GVPGRDGSFGANGIPGTPTGIPGRDGFKEGKEGECLERESFEBSWTPTYKQCSWSLNYGIDL 120
      |||
Qy      121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDO 180
      |||
Db      121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDO 180
      |||
Qy      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
      |||
Db      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
      |||
Qy      241 LPK 243
      |||
Db      241 LPK 243
```

```
      ; APPLICANT: Michael MORRISEY
      ; APPLICANT: Peter OLANDT
      ; APPLICANT: Ami SEN
      ; APPLICANT: Peter VEIBY
      ; APPLICANT: Gordon B. MILLS
      ; APPLICANT: Robert C. BAST, Jr.
      ; APPLICANT: Karen LU
      ; APPLICANT: Rosemarie SCHMANT
      ; APPLICANT: Xumei ZHAO
      ; APPLICANT: Karen GLATT
      ; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
      ; FILE REFERENCE: MRI-030
      ; CURRENT APPLICATION NUMBER: US/10/097,340
      ; CURRENT FILING DATE: 2002-03-14
      ; PRIOR APPLICATION NUMBER: 60/276,025
      ; PRIOR FILING DATE: 2001-03-14
      ; PRIOR APPLICATION NUMBER: 60/325,149
      ; PRIOR FILING DATE: 2001-09-26
      ; PRIOR APPLICATION NUMBER: 60/276,026
      ; PRIOR FILING DATE: 2001-03-14
      ; PRIOR APPLICATION NUMBER: 60/324,967
      ; PRIOR FILING DATE: 2001/09/26
      ; PRIOR APPLICATION NUMBER: 60/311,732
      ; PRIOR FILING DATE: 2001-08-10
      ; PRIOR APPLICATION NUMBER: 60/325,102
      ; PRIOR FILING DATE: 2001-09-26
      ; PRIOR APPLICATION NUMBER: 60/323,580
      ; PRIOR FILING DATE: 2001-09-19
      ; NUMBER OF SEQ ID NOS: 363
      ; SOFTWARE: FastSeq for Windows Version 4.0
      ; SEQ ID NO 135
      ; LENGTH: 243
      ; TYPE: PRT
      ; ORGANISM: Homo sapiens
US-10-097-340-135

Query Match      100.0%; Score 1303; DB 14; Length 243;
Best Local Similarity 100.0%; Pred. No.1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MRQGPASFORLRGLLLLLQLPAPSSASBIPKQKQAKQLRQREVVLDYNGMCLQCPA 60
      |||
Db      1 MRQGPASFORLRGLLLLLQLPAPSSASBIPKQKQAKQLRQREVVLDYNGMCLQCPA 60
      |||
Qy      61 GVPGRDGSFGANGIPGTPTGIPGRDGFKEGKEGECLERESFEBSWTPTYKQCSWSLNYGIDL 120
      |||
Db      61 GVPGRDGSFGANGIPGTPTGIPGRDGFKEGKEGECLERESFEBSWTPTYKQCSWSLNYGIDL 120
      |||
Qy      121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDO 180
      |||
Db      121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIITYLDO 180
      |||
Qy      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
      |||
Db      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
      |||
Qy      241 LPK 243
      |||
Db      241 LPK 243

RESULT 4
US-10-295-027-50
; Sequence 50, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
```

```

; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,249
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/356,714
; REMAINING PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 50
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-50
```

```

Query Match      100.0%; Score 1303; DB 15; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPQGPASPORLRLGILLILLQLPAPSSASIPKQKQAKLRQREVVDLYNGMCLQGPA 60
DB      1 MRPQGPASPORLRLGILLILLQLPAPSSASIPKQKQAKLRQREVVDLYNGMCLQGPA 60

QY      61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESFEEWTPNYKQCSMSLNYGIDL 120
DB      61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESFEEWTPNYKQCSMSLNYGIDL 120

QY      121 GKIACTFTKMSNSALRLVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLQ 180
DB      121 GKIACTFTKMSNSALRLVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLQ 180

QY      181 GSPENASTNINHRSTSVEGLCEGIGAGLVDAIWMGTGSDPYKGDASTGMSVSRITIEE 240
DB      181 GSPENASTNINHRSTSVEGLCEGIGAGLVDAIWMGTGSDPYKGDASTGMSVSRITIEE 240

QY      241 LPK 243
DB      241 LPK 243
```

```

RESULT 5
US-10-295-027-1173
; Sequence 1173, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Atiz, Natascha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
```

```

; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,249
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/356,714
; REMAINING PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 1173
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1173
```

```

Query Match      100.0%; Score 1303; DB 15; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPQGPASPORLRLGILLILLQLPAPSSASIPKQKQAKLRQREVVDLYNGMCLQGPA 60
DB      1 MRPQGPASPORLRLGILLILLQLPAPSSASIPKQKQAKLRQREVVDLYNGMCLQGPA 60

QY      61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESFEEWTPNYKQCSMSLNYGIDL 120
DB      61 GVPGRDGSFGANGIPGTGIPGRDGFKEGKGLRESFEEWTPNYKQCSMSLNYGIDL 120

QY      121 GKIACTFTKMSNSALRLVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLQ 180
DB      121 GKIACTFTKMSNSALRLVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLQ 180

QY      181 GSPENASTNINHRSTSVEGLCEGIGAGLVDAIWMGTGSDPYKGDASTGMSVSRITIEE 240
DB      181 GSPENASTNINHRSTSVEGLCEGIGAGLVDAIWMGTGSDPYKGDASTGMSVSRITIEE 240

QY      241 LPK 243
DB      241 LPK 243
```

```

RESULT 6
US-10-173-999-74
; Sequence 74, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
```

```
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; FILE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-74

Query Match          100.0%; Score 1303; DB 15; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPASPORLRGLLLLLQLPAPSSASIPKQKQAKQRLRQREVVLDVYNGMCLQGPA 60
    |||
DB 1 MRPQGPASPORLRGLLLLLQLPAPSSASIPKQKQAKQRLRQREVVLDVYNGMCLQGPA 60
QY 61 GVPGRDGPANGIPGTGIPGRDGFKEGKECLRESFEESWTPYKQCSMSLNYGIDL 120
    |||
DB 61 GVPGRDGPANGIPGTGIPGRDGFKEGKECLRESFEESWTPYKQCSMSLNYGIDL 120
QY 121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYDQ 180
    |||
DB 121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYDQ 180
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
    |||
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
QY 241 LPK 243
    |||
DB 241 LPK 243

RESULT 7
US-10-058-270A-38
; Sequence 38, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Afari, Daniel
; TITLE OF INVENTION: Methods of diagnosis of Breast Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer
; FILE REFERENCE: 018501-005210US
; CURRENT APPLICATION NUMBER: US/10/058,270A
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 60/263,965
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 60/265,928
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 09/829,472
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/282,698
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/288,590
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,443
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 141
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-270A-38

Query Match          100.0%; Score 1303; DB 15; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPASPORLRGLLLLLQLPAPSSASIPKQKQAKQRLRQREVVLDVYNGMCLQGPA 60
    |||
DB 1 MRPQGPASPORLRGLLLLLQLPAPSSASIPKQKQAKQRLRQREVVLDVYNGMCLQGPA 60
QY 61 GVPGRDGPANGIPGTGIPGRDGFKEGKECLRESFEESWTPYKQCSMSLNYGIDL 120
    |||
DB 61 GVPGRDGPANGIPGTGIPGRDGFKEGKECLRESFEESWTPYKQCSMSLNYGIDL 120
QY 121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYDQ 180
    |||
DB 121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYDQ 180
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
    |||
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240

RESULT 8
US-10-634-108-4
; Sequence 4, Application US/10634108
; Publication No. US20040063176A1
; GENERAL INFORMATION:
; APPLICANT: PRIESEL, Robert F.
; APPLICANT: LINDNER, Volkhart
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/10/634,108
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-108-4

Query Match          100.0%; Score 1303; DB 15; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPASPORLRGLLLLLQLPAPSSASIPKQKQAKQRLRQREVVLDVYNGMCLQGPA 60
    |||
DB 1 MRPQGPASPORLRGLLLLLQLPAPSSASIPKQKQAKQRLRQREVVLDVYNGMCLQGPA 60
QY 61 GVPGRDGPANGIPGTGIPGRDGFKEGKECLRESFEESWTPYKQCSMSLNYGIDL 120
    |||
DB 61 GVPGRDGPANGIPGTGIPGRDGFKEGKECLRESFEESWTPYKQCSMSLNYGIDL 120
QY 121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYDQ 180
    |||
DB 121 GKIACTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYDQ 180
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
    |||
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIEE 240
```



QY 241 LPK 243  
|||  
Db 241 LPK 243

## RESULT 9

US-10-188-832-175  
; Sequence 175, Application US/10188832  
; Publication No. US20040076955A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David H.  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions  
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 018501-002330US  
; CURRENT APPLICATION NUMBER: US/10/188,832  
; CURRENT FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: US 60/302,814  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/310,099  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: US 60/343,705  
; PRIOR FILING DATE: 2001-11-08  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/372,246  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 175  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-832-175

Query Match 100.0%; Score 1303; DB 15; Length 243;  
Best Local Similarity 100.0%; Pred. No. 1.5e-115;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNMCLOGPA 60  
Db 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNMCLOGPA 60  
QY 61 GVPGRDGSFGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPNTYKQCSWSSLNYGIDL 120  
Db 61 GVPGRDGSFGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPNTYKQCSWSSLNYGIDL 120  
QY 121 GKIAECTFTKMSNSALRYLFGSGLRLKCRNACCCQWYTFNGAECSSGPLPIAIIYLDQ 180  
Db 121 GKIAECTFTKMSNSALRYLFGSGLRLKCRNACCCQWYTFNGAECSSGPLPIAIIYLDQ 180  
QY 181 GSPENNSTINIHRTSVEGICGEGIGAGLVDAIIVGTGSDYPKGDASTGNSVSRITIEE 240  
Db 181 GSPENNSTINIHRTSVEGICGEGIGAGLVDAIIVGTGSDYPKGDASTGNSVSRITIEE 240  
QY 241 LPK 243  
|||  
Db 241 LPK 243

## RESULT 10

US-10-939-233-4  
; Sequence 4, Application US/10939233  
; Publication No. US20050147602A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volhard  
; APPLICANT: PRIESEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTRHCL, A NOVEL,  
; FILE REFERENCE: 053689-5006-03  
; CURRENT APPLICATION NUMBER: US/10/939,233

; CURRENT FILING DATE: 2004-09-10  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US 10/045,992  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-939-233-4

Query Match 100.0%; Score 1303; DB 18; Length 243;  
Best Local Similarity 100.0%; Pred. No. 1.5e-115;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNMCLOGPA 60  
Db 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQKQALRQREVVDLYNMCLOGPA 60  
QY 61 GVPGRDGSFGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPNTYKQCSWSSLNYGIDL 120  
Db 61 GVPGRDGSFGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPNTYKQCSWSSLNYGIDL 120  
QY 121 GKIAECTFTKMSNSALRYLFGSGLRLKCRNACCCQWYTFNGAECSSGPLPIAIIYLDQ 180  
Db 121 GKIAECTFTKMSNSALRYLFGSGLRLKCRNACCCQWYTFNGAECSSGPLPIAIIYLDQ 180  
QY 181 GSPENNSTINIHRTSVEGICGEGIGAGLVDAIIVGTGSDYPKGDASTGNSVSRITIEE 240  
Db 181 GSPENNSTINIHRTSVEGICGEGIGAGLVDAIIVGTGSDYPKGDASTGNSVSRITIEE 240  
QY 241 LPK 243  
|||  
Db 241 LPK 243

## RESULT 11

US-10-097-340-131  
; Sequence 131, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVARAPU  
; APPLICANT: Sebastian HOERCH  
; APPLICANT: Shubhangi KAMATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
; FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; CURRENT FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732

;; PRIOR FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: 60/325,102  
;; PRIOR FILING DATE: 2001-09-26  
;; PRIOR APPLICATION NUMBER: 60/323,580  
;; PRIOR FILING DATE: 2001-09-19  
;; NUMBER OF SEQ ID NOS: 363  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 131  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-097-340-131

Query Match Best Local Similarity 100.0%; Score 1303; DB 14; Length 278;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRGLLLLLQLPAPSSASEIPKQKQAKLRQREVVDLYNGMCLQSPA 60  
DB 36 MRPGPAASPORLRGLLLLLQLPAPSSASEIPKQKQAKLRQREVVDLYNGMCLQSPA 95  
QY 61 GVPGRDGSFGANGIPGTGIGIPGRDGFKEGKEGCLRESFEESWTPTYKQCSMSLNYGIDL 120  
DB 96 GVPGRDGSFGANGIPGTGIGIPGRDGFKEGKEGCLRESFEESWTPTYKQCSMSLNYGIDL 155  
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCCQWFTFNGACSGPLPEAIITYLDQ 180  
DB 156 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCCQWFTFNGACSGPLPEAIITYLDQ 215  
QY 181 GSPENNSTINIHRTSVSVEGLCEGIGAGLVDAIIVGTCSDDYPKGASTGMNSVSR11IEE 240  
DB 216 GSPENNSTINIHRTSVSVEGLCEGIGAGLVDAIIVGTCSDDYPKGASTGMNSVSR11IEE 275  
QY 241 LPK 243  
DB 276 LPK 278

## RESULT 12

US-10-177-293-496  
;; Sequence 496, Application US/10177293  
;; Publication No. US20030124128A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lillie, James  
;; APPLICANT: Glatz, Karen  
;; APPLICANT: Zhao, Xumei  
;; APPLICANT: Gannavarpu, Manjula  
;; APPLICANT: Kamatkar, Shubhangi  
;; APPLICANT: Mertens, Maureen  
;; APPLICANT: Myer, Vic  
;; APPLICANT: Wang, Youzhen  
;; APPLICANT: Xu, Yongyao  
;; APPLICANT: Hoersch, Sebastien  
;; APPLICANT: Monahan, John  
;; APPLICANT: Meyers, Rachel E.  
;; APPLICANT: Bast Jr., Robert C.  
;; APPLICANT: Hortobagyi, Gabriel N.  
;; APPLICANT: Pusztai, Lajos  
;; APPLICANT: Meric, Funda  
;; APPLICANT: Sahin, Aysegül  
;; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,  
;; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER  
;; FILE REFERENCE: MRI-038  
;; CURRENT APPLICATION NUMBER: US/10/177,293  
;; PRIOR FILING DATE: 2002-06-21  
;; PRIOR APPLICATION NUMBER: US 60/299,887  
;; PRIOR FILING DATE: 2001-06-21  
;; PRIOR APPLICATION NUMBER: US 60/301,572  
;; PRIOR FILING DATE: 2001-06-27  
;; PRIOR APPLICATION NUMBER: US 60/306,501  
;; PRIOR FILING DATE: 2001-07-18  
;; PRIOR APPLICATION NUMBER: US 60/325,002

;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: US 60/362,585  
;; PRIOR FILING DATE: 2002-03-05  
;; PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
;; PRIOR FILING DATE: 2002-05-14  
;; NUMBER OF SEQ ID NOS: 506  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 496  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-177-293-496

Query Match Best Local Similarity 100.0%; Score 1303; DB 14; Length 278;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRGLLLLLQLPAPSSASEIPKQKQAKLRQREVVDLYNGMCLQSPA 60  
DB 36 MRPGPAASPORLRGLLLLLQLPAPSSASEIPKQKQAKLRQREVVDLYNGMCLQSPA 95  
QY 61 GVPGRDGSFGANGIPGTGIGIPGRDGFKEGKEGCLRESFEESWTPTYKQCSMSLNYGIDL 120  
DB 96 GVPGRDGSFGANGIPGTGIGIPGRDGFKEGKEGCLRESFEESWTPTYKQCSMSLNYGIDL 155  
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCCQWFTFNGACSGPLPEAIITYLDQ 180  
DB 156 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCCQWFTFNGACSGPLPEAIITYLDQ 215  
QY 181 GSPENNSTINIHRTSVSVEGLCEGIGAGLVDAIIVGTCSDDYPKGASTGMNSVSR11IEE 240  
DB 216 GSPENNSTINIHRTSVSVEGLCEGIGAGLVDAIIVGTCSDDYPKGASTGMNSVSR11IEE 275  
QY 241 LPK 243  
DB 276 LPK 278

## RESULT 13

US-10-301-822-77  
;; Sequence 77, Application US/10301822  
;; Publication No. US20030148410A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Millennium Pharmaceuticals, Inc.  
;; APPLICANT: Berger, Allison  
;; APPLICANT: Guillemette, Tracy L.  
;; APPLICANT: Kamatkar, Shubhangi  
;; APPLICANT: Schlegel, Robert  
;; APPLICANT: Monahan, John E.  
;; APPLICANT: Thibodeau, Stephen N.  
;; APPLICANT: Burgart, Lawrence J.  
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
;; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
;; TITLE OF INVENTION: THERAPY OF COLON CANCER  
;; FILE REFERENCE: MP001-029P2RNM  
;; CURRENT APPLICATION NUMBER: US/10/301,822  
;; PRIOR FILING DATE: 2002-11-21  
;; PRIOR APPLICATION NUMBER: US 60/339,971  
;; PRIOR FILING DATE: 2001-12-10  
;; PRIOR APPLICATION NUMBER: US 60/361,978  
;; PRIOR FILING DATE: 2002-03-05  
;; PRIOR APPLICATION NUMBER: US 60/381,988  
;; PRIOR FILING DATE: 2002-05-20  
;; NUMBER OF SEQ ID NOS: 228  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 77  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-301-822-77

Query Match Best Local Similarity 100.0%; Score 1303; DB 14; Length 278;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNGMCLQGPA 60  
 DB 36 MRPGPASPQRRLGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNGMCLQGPA 95

QY 61 GVGGRDGPANGIPGTGGIPGRDGFGEKGECLRESFEESWTPNYKQCSWSSLYNGIDL 120  
 DB 96 GVGGRDGPANGIPGTGGIPGRDGFGEKGECLRESFEESWTPNYKQCSWSSLYNGIDL 155

QY 121 GKIAECTFTKRSNSALRVLPFSGSLRLKCRNACQORWYTFNGAECGSLPIEAIYYLDQ 180  
 DB 156 GKIAECTFTKRSNSALRVLPFSGSLRLKCRNACQORWYTFNGAECGSLPIEAIYYLDQ 215

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRIIIEE 240  
 DB 216 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRIIIEE 275

QY 241 LPK 243  
 DB 276 LPK 278

RESULT 14

US-10-296-115-1261  
 ; Sequence 1261, Application US/10296115  
 ; Publication No. US20040053248A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyeg Inc  
 ; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides  
 ; FILE REFERENCE: 784PCT  
 ; CURRENT APPLICATION NUMBER: US/10/296,115  
 ; PRIOR FILING DATE: 2002-11-18  
 ; PRIOR APPLICATION NUMBER: US09/488,725  
 ; PRIOR FILING DATE: 2000-01-21  
 ; PRIOR APPLICATION NUMBER: US09/552,317  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 1478  
 ; SEQ ID NO 1261  
 ; LENGTH: 278  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-296-115-1261

Query Match 100.0%; Score 1303; DB 15; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-115;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNGMCLQGPA 60  
 DB 36 MRPGPASPQRRLGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNGMCLQGPA 95

QY 61 GVGGRDGPANGIPGTGGIPGRDGFGEKGECLRESFEESWTPNYKQCSWSSLYNGIDL 120  
 DB 96 GVGGRDGPANGIPGTGGIPGRDGFGEKGECLRESFEESWTPNYKQCSWSSLYNGIDL 155

QY 121 GKIAECTFTKRSNSALRVLPFSGSLRLKCRNACQORWYTFNGAECGSLPIEAIYYLDQ 180  
 DB 156 GKIAECTFTKRSNSALRVLPFSGSLRLKCRNACQORWYTFNGAECGSLPIEAIYYLDQ 215

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRIIIEE 240  
 DB 216 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRIIIEE 275

QY 241 LPK 243  
 DB 276 LPK 278

RESULT 15  
 US-10-961-139-2  
 ; Sequence 2, Application US/10961139  
 ; Publication No. US2005015313A1

GENERAL INFORMATION:  
 ; APPLICANT: Endege, Wilson O.  
 ; APPLICANT: Ford, Donna  
 ; APPLICANT: Ganavaru, Manjula  
 ; APPLICANT: Glat, Karen  
 ; APPLICANT: Hoersch, Sebastian  
 ; APPLICANT: Kamakar, Shubhangi  
 ; APPLICANT: Monahan, John E.  
 ; APPLICANT: Schlegel, Robert  
 ; APPLICANT: Xu, Yong Yao  
 ; APPLICANT: Zhao, Xumei  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND PROTEINS FOR  
 ; TITLE OF INVENTION: THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
 ; FILE REFERENCE: MRI-065  
 ; CURRENT APPLICATION NUMBER: US/10/961,139  
 ; PRIOR FILING DATE: 2004-10-07  
 ; PRIOR APPLICATION NUMBER: 60/509,171  
 ; PRIOR FILING DATE: 2003-10-07  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 278  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-10-961-139-2

Query Match 100.0%; Score 1303; DB 18; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-115;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNGMCLQGPA 60  
 DB 36 MRPGPASPQRRLGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNGMCLQGPA 95

QY 61 GVGGRDGPANGIPGTGGIPGRDGFGEKGECLRESFEESWTPNYKQCSWSSLYNGIDL 120  
 DB 96 GVGGRDGPANGIPGTGGIPGRDGFGEKGECLRESFEESWTPNYKQCSWSSLYNGIDL 155

QY 121 GKIAECTFTKRSNSALRVLPFSGSLRLKCRNACQORWYTFNGAECGSLPIEAIYYLDQ 180  
 DB 156 GKIAECTFTKRSNSALRVLPFSGSLRLKCRNACQORWYTFNGAECGSLPIEAIYYLDQ 215

QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRIIIEE 240  
 DB 216 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRIIIEE 275

QY 241 LPK 243  
 DB 276 LPK 278

Search completed: August 4, 2005, 04:45:27  
 Job time : 87.0824 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:56 ; Search time 24.6601 Seconds  
(without alignments) 741.643 Million cell updates/sec

Title: US-10-634-108-2

Perfect score: 1307  
Sequence: 1 MHQGRASPQLLGLFLV.....GDASTGMSVSRITIELPK 245

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5A COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6A COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PTUS COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1307	100.0	245	4	US-09-692-081-2 Sequence 2, Appl1
2	1307	100.0	277	4	US-09-692-081-5 Sequence 5, Appl1
3	1225	93.7	243	4	US-09-692-081-4 Sequence 4, Appl1
4	1222	93.5	243	4	US-09-489-847-205 Sequence 205, App
5	1216	93.0	243	4	US-09-834-759-514 Sequence 514, App
6	1216	93.0	278	4	US-09-834-759-515 Sequence 515, App
7	1101	84.2	276	4	US-09-205-258-958 Sequence 958, App
8	968	74.1	197	4	US-09-834-759-516 Sequence 516, App
9	968	74.1	232	4	US-09-834-759-517 Sequence 517, App
10	353	27.0	66	4	US-09-205-258-962 Sequence 962, App
11	296	22.6	52	4	US-09-205-258-961 Sequence 961, App
12	270	20.7	51	4	US-09-205-258-963 Sequence 963, App
13	248	19.0	46	4	US-09-834-759-518 Sequence 518, App
14	215	16.4	52	4	US-09-205-258-960 Sequence 960, App
15	141	10.8	26	4	US-09-834-759-519 Sequence 519, App
16	130	9.9	93	4	US-09-489-847-240 Sequence 240, App
17	130	9.9	93	4	US-09-489-847-362 Sequence 362, App
18	124.5	9.5	1516	4	US-09-949-016-8209 Sequence 8209, App
19	121	9.3	128	3	US-09-227-357-190 Sequence 190, App
20	121	9.3	623	3	US-09-029-348-3 Sequence 3, Appl1
21	120	9.2	626	3	US-09-029-348-2 Sequence 2, Appl1
22	119.5	9.1	755	4	US-09-919-497-57 Sequence 57, Appl1
23	119.5	9.1	246	2	US-08-463-911-4 Sequence 4, Appl1
24	119.5	9.1	313	4	US-09-949-016-9265 Sequence 9265, App
25	117	9.0	1670	4	US-09-949-016-5883 Sequence 5883, App
26	117	9.0	1057	3	US-08-931-820-4 Sequence 4, Appl1
27	117	9.0	1078	3	US-08-963-825-21 Sequence 21, Appl1

28	117	9.0	1078	3	US-09-500-811-21 Sequence 21, Appl1
29	117	9.0	1078	3	US-09-570-573-21 Sequence 21, Appl1
30	117	9.0	1078	3	US-09-548-608-21 Sequence 21, Appl1
31	116	8.9	258	4	US-09-976-594-815 Sequence 815, App
32	116	8.9	744	4	US-09-949-016-9607 Sequence 9607, App
33	115.5	8.8	349	6	5510466-2 Patent No. 5510466
34	115.5	8.8	349	6	5510466-2 Patent No. 5510466
35	115.5	8.8	453	6	5510466-4 Patent No. 5510466
36	115.5	8.8	453	6	5510466-4 Patent No. 5510466
37	115.5	8.8	684	3	US-08-555-669-12 Sequence 12, Appl1
38	115.5	8.8	684	3	US-09-073-663-12 Sequence 12, Appl1
39	115	8.8	245	4	US-09-552-225A-4 Sequence 4, Appl1
40	115	8.8	245	4	US-09-552-204A-4 Sequence 4, Appl1
41	115	8.8	1078	4	US-09-949-016-11185 Sequence 11185, App
42	115	8.8	1143	4	US-09-949-016-6137 Sequence 6137, App
43	114.5	8.8	248	4	US-09-600-932-28 Sequence 28, Appl1
44	114	8.7	310	3	US-09-219-849-47 Sequence 47, Appl1
45	114	8.7	451	1	US-08-453-117-2 Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-692-081-2
; Sequence 2, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: PRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-692-081-2

Query Match      100.0%; Score 1307; DB 4; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.2e-133; Indels 0; Gaps 0;
Matches 245; Conservative 0; Mismatches 0;

QY      1 MHQGRASPQLLGLFLVLLQLSAPSSASBNPKVKOKALIROEVVDLYNGMCLOG 60
      1 MHQGRASPQLLGLFLVLLQLSAPSSASBNPKVKOKALIROEVVDLYNGMCLOG 60
Db
QY      61 PAGVPRDGSPPANGIPGTGIPGRDGFKEGKECELRSEFESWTPNPKQCSWSSLNYGI 120
      61 PAGVPRDGSPPANGIPGTGIPGRDGFKEGKECELRSEFESWTPNPKQCSWSSLNYGI 120
Db
QY      121 DGLKIAECFTMRNSALRVLFSGSLRKLKCNACCORRYFTFNAGECGPIPIEATITVL 180
      121 DGLKIAECFTMRNSALRVLFSGSLRKLKCNACCORRYFTFNAGECGPIPIEATITVL 180
Db
QY      181 DQGSBELNSTINIHRTSSVEGLCEGIGAGLVVAIVWGTCSDPKGDASTGMNSVSRIT 240
      181 DQGSBELNSTINIHRTSSVEGLCEGIGAGLVVAIVWGTCSDPKGDASTGMNSVSRIT 240
Db
QY      241 BELPK 245
      241 BELPK 245
Db
QY      241 BELPK 245
      241 BELPK 245
Db

RESULT 2
US-09-692-081-5
; Sequence 5, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
```

```

: APPLICANT: FRIESEL, Robert F.
: TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
: FILE REFERENCE: 36-13 Linder et al. (200036.0013)
: CURRENT APPLICATION NUMBER: US/09/692,081
: CURRENT FILING DATE: 2000-10-19
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 5
: LENGTH: 277
: TYPE: PRT
: ORGANISM: Rattus sp.
: US-09-692-081-5

```

Query Match	100.0%;	Score 1307;	DB 4;	Length 277;
Best Local Similarity	100.0%;	Pred. No. 2.6e-133;		
Matches 245;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MHPQGRASPOLLLGFLVLLVLLLLQLSAPSSASNPVKQKALLRQRPVDVLYNGMCLQG	60
Db	33	MHPQGRASPOLLLGFLVLLVLLLLQLSAPSSASNPVKQKALLRQRPVDVLYNGMCLQG	92
Qy	61	PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGDECLARESPESWTPNNVKQCSWSLNYGI	120
Db	93	PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGDECLARESPESWTPNNVKQCSWSLNYGI	152
Qy	121	DLGKIACCTTKRNSALRVLPFSGSLRLKCNACCCORWYTFENGABCSGFLPEALITYL	180
Db	153	DLKIAECTTKRNSALRVLPFSGSLRLKCNACCCORWYTFENGABCSGFLPEALITYL	212
Qy	181	DQGSPELNTSTINIHRTSSVEGLCEGIGAGLVVVALVWNGCSDPYKGDASTOMNSVSRIT	240
Db	213	DQGSPELNTSTINIHRTSSVEGLCEGIGAGLVVVALVWNGCSDPYKGDASTOMNSVSRIT	272
Qy	241	EELLPK 245	
Db	273	EELLPK 277	

```

RESULT 3
US-09-692-081-4
; Sequence 4, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDERER, Volhard
; APPLICANT: FRIESEB, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS
; FILE REFERENCE: 36-13 Linderer et al. (2000036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; CURRENT FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;
; LENGTH: 243
;
; TYPE: PR1
;
; ORGANISM: Homo sapiens
US-09-692-081-4

```

[illegible]

Qy	181	DOSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTMNSVSR	11	240
Db	179	DQSPPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTMNSVSR	11	238
Qy	241	EELPK	245	
Db	239	EELPK	243	

QY	241	EELPK	245
Db	239	EELPK	243

```

RESULT 4
US-09-489-847-205
; Sequence 205, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031p1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 205
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-205

```

```

Query Match: 93.5%; Score 1222; DB 4; Length 243;
Best Local Similarity 94.3%; Pred. No. 3,6e-124;
Matches 231; Conservative 4; Mismatches 8; Indels 2; Gaps 1

QY      1 MHFGAASPOLLLGLFVLLLLQLSAPSSASBNPKOKALLROREVDVLYNGMCLQG 60
Db      1 MRQGPAAAPQRLRG--LLLLLLLLQLPAPSSASEIPKQKQAKQRLQREVDVLYNGMCLQG 58

QY      61 PAGVPPRDSPGANGIPGTPGIPGPDGFKGEGECLRSFEESWTPNYKQCSMSLNYGI 120
Db      59 PAGVPPRDSPGANGIPGTPGIPGPDGFKGEGECLRSFEESWTPNYKQCSMSLNYGI 118

QY      121 DLGKIACEFTKQRSNALRYLFGSLLRLKCNACCOMWYFTPNGAECSGPLPEAIYYL 180
Db      119 DLGKIACEFTKQRSNALRYLFGSLLRLKCNACCEMYFTPNGAECSGPLPEAIYYL 178

QY      181 DQSPPELNTINIHRTSSVEGLCEGIGAGLYDVVAIWTGCSDPYKGDASTGMSVSR11 240
Db      179 DQSPPELNTINIHRTSSVEGLCEGIGAGLYDVVAIWTGCSDPYKGDASTGMSVSR11 238

QY      241 EELPK 245
Db      239 EELPK 243

RESULT 5
US-09-834-759-514
; Sequence 514, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun

```

```
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ. ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 514
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-514

Query Match      93.0%; Score 1216; DB 4; Length 243;
Best Local Similarity 94.3%; Pred. No. 1.6e-123;
Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY      1 MHPOGRASPOLLLGLFVLLILLQLSAPSSASENPVKOKALIRQREVVDLYNGMCTLG 60
      1 MRQGPAPASPORLGG--LTLILLLLQLPAPSSASEIPKQKQALRQREVVDLYNGMCTLG 58
Db      59 PAVPGRDSSPGANGIPGTGPIGRDGFKEGKEGECLEESPEESWTPNYKOCSSSLNYGI 120
      121 DLGKIACFTFTKRRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYL 180
      119 DLGKIACFTFTKRRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYL 178
QY      181 DQSSPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIII 240
      179 DQSSPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIII 238
Db      241 EELPK 245
      239 EELPK 243

RESULT 6
US-09-834-759-515
; Sequence 515, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-515

Query Match      93.0%; Score 1216; DB 4; Length 278;
Best Local Similarity 94.3%; Pred. No. 2e-123;
Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY      1 MHPOGRASPOLLLGLFVLLILLQLSAPSSASENPVKOKALIRQREVVDLYNGMCTLG 60
      36 MRQGPAPASPORLGG--LTLILLLLQLPAPSSASEIPKQKQALRQREVVDLYNGMCTLG 93
Db
```

```
QY      61 PAVPGRDSSPGANGIPGTGPIGRDGFKEGKEGECLEESPEESWTPNYKOCSSSLNYGI 120
      94 PAVPGRDSSPGANVLPGTGPIGRDGFKEGKEGECLEESPEESWTPNYKOCSSSLNYGI 153
Db      121 DLGKIACFTFTKRRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYL 180
      154 DLGKIACFTFTKRRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYL 213
QY      181 DQSSPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIII 240
      214 DQSSPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIII 273
Db      241 EELPK 245
      274 EELPK 278

RESULT 7
US-09-205-258-958
; Sequence 958, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 958
; LENGTH: 276
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-958

```

```

Query Match      84.2%; Score 1101; DB 4; Length 276;
Best Local Similarity 98.5%; Pred. No. 5.7e-111;
Matches 202; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 41 KALIRREVVLDLYNGMCLCGPAGVPGRDGSPGANGIPGTGIGRGGFKGKCECLRESE 100
DB 72 KRIIRREVVLDLYNGMCLCGPAGVPGRDGSPGANGIPGTGIGRGGFKGKCECLRESE 131
QY 101 EESWTPNYKCCSSSLNYGIDLCIAECTTKKRSNSALRVLPSSGLRLKCRNACCOMRY 160
DB 132 EESWTPNYKCCSSSLNYGIDLCIAECTTKKRSNSALRVLPSSGLRLKCRNACCOMRY 191
QY 161 FTNNGACSGPLPIEAIIVYDQGSPELNTINIHRTSSVGLCEGIGAGIVDAIVWGTC 220
DB 192 FTNNGACSGPLPIEAIIVYDQGSPELNTINIHRTSSVGLCEGIGAGIVDAIVWGTC 251
QY 221 SDYKGDASTGMSVSRIITIEELPK 245
DB 252 SDYKGDASTGMSVSRIITIEELPK 276

```

```

RESULT 8
US-09-834-759-516
; Sequence 516, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:

```

```

; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 516
; LENGTH: 197
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-834-759-516

```

```

Query Match      74.1%; Score 968; DB 4; Length 197;
Best Local Similarity 93.0%; Pred. No. 9.3e-97;
Matches 185; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

```

```

QY 1 MHPGRRASPQULLGLFVLVLLQLSAPSSASENPKYKQKALIRQREVVDLYNGMCLCG 60
DB 1 MHPGRRASPQRLRG--LILLILLQLPAPSSASEIPKQKQKQRLQREVVDLYNGMCLCG 58
QY 61 PAGVPGRDGSPGANGIPGTGIGRGGFKGKCECLRESESWTPNYKCCSSSLNYGI 120
DB 59 PAGVPGRDGSPGANVYPGTGGIGRGGFKGKCECLRESESWTPNYKCCSSSLNYGI 118
QY 121 DLGIAECTFTKRSNSALRVLPSSGLRLKCRNACCOMRYFTNNGACSGPLPIEAIIVL 180
DB 119 DLGIAECTFTKRSNSALRVLPSSGLRLKCRNACCOMRYFTNNGACSGPLPIEAIIVL 178
QY 181 DQGSPELNTINIHRTSSV 199
DB 179 DQGSPELNTINIHRTSSV 197

```

```

RESULT 9
US-09-834-759-517
; Sequence 517, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 517
; LENGTH: 232
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-834-759-517

```

```

Query Match      74.1%; Score 968; DB 4; Length 232;
Best Local Similarity 93.0%; Pred. No. 1.2e-96;
Matches 185; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

```

```

QY 1 MHPGRRASPQULLGLFVLVLLQLSAPSSASENPKYKQKALIRQREVVDLYNGMCLCG 60
DB 36 MHPGRRASPQRLRG--LILLILLQLPAPSSASEIPKQKQKQRLQREVVDLYNGMCLCG 93

```



```
QY 61 PAGVGRDGSBANGIPGTBGPGRDGFKEGKGBCLRESFEESGTPYKOCSSSLNYGI 120
DB 94 PAGVGRDGSBANGIPGTBGPGRDGFKEGKGBCLRESFEESGTPYKOCSSSLNYGI 153
QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYL 180
DB 154 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYL 213
QY 181 DOGSPELNTINIRITSSV 199
DB 214 DOGSPELNTINIRITSSV 232

RESULT 10
US-09-205-258-962
; Sequence 962, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER APPLICATION NUMBER: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
```

```
EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 962
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-962

Query Match 27.0%; Score 353; DB 4; Length 66;
Best Local Similarity 98.5%; Pred. No. 9,3e-31;
Matches 65; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 130 FTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYLDOGSPELNS 189
DB 1 FTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYYLDOGSPELNS 60
QY 190 TINIR 195
DB 61 TINIR 66

RESULT 11
US-09-205-258-961
; Sequence 961, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER APPLICATION NUMBER: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 961
; LENGTH: 52
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-205-258-961
;
; Query Match 22.6%; Score 296; DB 4; Length 52;
; Best Local Similarity 100.0%; Pred. No. 1e-24;
; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Db 79 TCGTGTGGGFKGKGCRCRESFEESWTPYKCCSSLNATGIDLGIACCTF 130
; 1 TGGTGTGGGFKGKGCRCRESFEESWTPYKCCSSLNATGIDLGIACCTF 52
;
; RESULT 12
; US-09-205-258-963
; Sequence 963, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
```

```

; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1327
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 963
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-963

```

```

Query Match      20.7%; Score 270; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 6,4e-22;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      195 RTSSVGLCGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIELPK 245
Db      1 RTSSVGLCGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIELPK 51

```

```

RESULT 13
US-09-834-759-518
; Sequence 518, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 518
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-518

```

```

Query Match      19.0%; Score 248; DB 4; Length 46;

```

```

Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      200 EGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIELPK 245
Db      1 EGLCEGIGAGLVDAIWTGTCSDYPKGDASTGMNSVSRITIELPK 46

```

```

RESULT 14
US-09-205-258-960
; Sequence 960, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917

```

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 960  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-205-258-960

Query Match 16.4%; Score 215; DB 4; Length 52;  
Best Local Similarity 95.2%; Pred. No. 6.1e-16;  
Matches 40; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 41 KALIRORVVDLYNGMCLOGPAGVPGRDSPGANGIPGTPGI 82  
DB 11 KRIIRORVVDLYNGMCLOGPAGVPGRDSPGANGIPGTPGI 52

RESULT 15  
US-09-834-759-519  
Sequence 519, Application US/09834759  
Patent No. 6680197  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yugu  
APPLICANT: Dillon, Devin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.470C9  
CURRENT APPLICATION NUMBER: US/09/834,759  
CURRENT FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 547  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 519  
LENGTH: 26  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-834-759-519

Query Match 10.8%; Score 141; DB 4; Length 26;  
Best Local Similarity 100.0%; Pred. No. 2.4e-08;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CSDYPKGDASTGMNSVSRRIIEELPK 245  
|||||

DB 1 CSDYPKGDASTGMNSVSRRIIEELPK 26  
Search completed: August 4, 2005, 04:23:52  
Job time : 25.6601 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:16 ; Search time 19.5359 Seconds

(without alignments)  
1206.653 Million cell updates/sec

Title: US-10-634-108-2

Perfect score: 1307

Sequence: 1 MHQGRAPQLLGFLVL.....GDASTGMSVSRRIIEELPK 245

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79: \*  
2: PIR 79: \*  
3: PIR 79: \*  
4: PIR 79: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	127.5	9.8	289	2	T20177	hypothetical prote
2	125.5	9.6	298	2	T27644	hypothetical prote
3	125.5	9.6	1752	2	A45407	collagen alpha 3(I
4	123.5	9.4	311	2	T15268	hypothetical prote
5	122.5	9.4	920	2	A45748	collagen alpha 1(V
6	122.5	9.4	1549	2	T48103	type VII collagen
7	122	9.3	330	2	S46557	collagen alpha 1(X
8	121	9.3	178	2	A39762	collagen alpha 1(X
9	121	9.3	1466	1	CGHU7L	collagen alpha 1(I
10	121	9.3	1763	1	S16366	collagen alpha 2(I
11	119.5	9.1	246	2	S29328	complement subcomp
12	119.5	9.1	310	2	T29731	hypothetical prote
13	119.5	9.1	323	2	A61396	collagen alpha 1(I
14	119.5	9.1	1670	1	CGHU3B	collagen alpha 3(I
15	119.5	9.1	2944	2	A54849	collagen alpha 1(V
16	119	9.1	298	2	JC1448	collagen col-34 -
17	119	9.1	299	2	T29956	hypothetical prote
18	119	9.1	1049	1	CGBO7S	collagen alpha 1(I
19	118	9.0	458	2	T31631	hypothetical prote
20	118	9.0	1669	1	CGMS4B	collagen alpha 1(I
21	118	9.0	1744	2	S41920	collagen alpha 1(I
22	117.5	9.0	324	2	A31920	collagen sct-1 pre
23	117.5	9.0	673	1	CGBO6C	collagen alpha 1(I
24	117	9.0	253	2	S49158	complement protein
25	116.5	8.9	58	2	A45526	collagen homolog A
26	116.5	8.9	300	2	T24482	hypothetical prote
27	116	8.9	289	2	T26812	hypothetical prote
28	116	8.9	302	2	T15936	hypothetical prote
29	116	8.9	744	2	S15435	collagen alpha 1(V

30	115.5	8.8	349	2	S08278	macrophage scaveng
31	115.5	8.8	428	2	T24769	hypothetical prote
32	115.5	8.8	453	2	S08276	macrophage scaveng
33	115	8.8	245	1	CHUQC	complement subcomp
34	115	8.8	296	2	A31219	collagen 1 - Caeno
35	115	8.8	301	2	B31219	collagen 2 - Caeno
36	115	8.8	744	1	S34246	collagen alpha 1(V
37	115	8.8	744	1	S23298	collagen alpha 1(V
38	115	8.8	1042	1	CGCH1S	collagen alpha 1(I
39	115	8.8	1142	2	JX0369	collagen alpha 1(X
40	114.5	8.8	248	1	LNHP1	pulmonary surfacta
41	114.5	8.8	1496	1	CGHU2V	collagen alpha 2(V
42	114.5	8.8	1497	2	I49607	procollagen type V
43	114	8.7	327	2	T29031	hypothetical prote
44	114	8.7	451	2	A38415	macrophage scaveng
45	114	8.7	671	1	CGRT1S	collagen alpha 1(I

## ALIGNMENTS

## RESULT 1

hypothetical protein C53B4.5 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #ext\_change 09-Jul-2004

C/Accession: T20177

R/Berke, M.

submitted to the EMBL Data Library, December 1995

A/Reference number: Z19233

A/Accession: T20177

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-289 <MIL>

A/Cross-references: UNIPROT:Q18799; EMBL:Z68215; PIDN:CAA92453.1; GSPDB:GN00022; CESP:C.

A/Experimental source: clone C53B4

C/Genetics:

A/Genes: CESP:C53B4.5

A/Map position: 4

Query Match

Best Local Similarity 9.8%; Score 127.5; DB 2; Length 289;

Matches 25; Conservative 4; Mismatches 8; Indels 15; Gaps 1;

QY

59 QCPAGVPRDGS-----PGANGTPTGTPGPRDGFGEKGC 95

DB 220 RGPAGQPGKDAQCGPGEKANGBPQGRDQPGRPQGRDGHGEKGC 271

RESULT 2

T27644 hypothetical protein ZK1010.7 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #ext\_change 09-Jul-2004

C/Accession: T27644

R/Gardner, A.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z20398

A/Accession: T27644

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-298 <MIL>

A/Cross-references: UNIPROT:O18286; EMBL:Z62083; PIDN:CAB04973.1; GSPDB:GN00021; CESP:ZK

A/Experimental source: clone ZK1010

C/Genetics:

A/Genes: CESP:ZK1010.7

A/Map position: 3

A/Introns: 54/3; 236/1

Query Match

Best Local Similarity 9.6%; Score 125.5; DB 2; Length 298;

Matches 30; Conservative 6; Mismatches 29; Indels 11; Gaps 2;

QY 31 SASENPXYKOKKLLIIROREVVDLYNGMCLGPGVGPGRDSSPGANGIIPGHPGIPGRGPGFG 90  
 Db 82 SSNETTRLLRLRDAAAC-----GCCLL---PGLPGDPGPGANGAPGRGAPGAPGPGF 130  
 QY 91 EKGECLRESFEESSWTP 106  
 Db 131 RPPAVCEEITPEPCTP 146

### RESULT 3

collagen alpha 1(V) chain - sea urchin (Strongylocentrotus purpuratus)  
C.Species: Strongylocentrotus purpuratus (purple urchin)  
C.Date: 22-Sep-1993 #sequence revision 18-Nov-1994 #text\_change 09-Jul-2004  
R.Expossto, J.Y.; D'Alesio, M.; Di Liberto, M.; Ramirez, F.  
J. Biol. Chem. 268, 5249-5254, 1993  
J.Title: Complete primary structure of a sea urchin type IV collagen alpha chain and ana  
A.Reference number: A45407; PMID:8444899  
A.Accession: A45407  
A.Status: preliminary; not compared with conceptual translation  
A.Molecule type: nucleic acid  
A.Residues: 1-1752 <EXP>  
A.Cross-references: UNIPROT:Q26312  
A.Note: sequence extracted from NCBI backbone (NCBIP:126841)  
R.Messel, G.M.; Ecklin, M.; Benson, S.  
Dev. Biol. 148, 261-272, 1991  
A.Title: Primary mesenchyme cells of the sea urchin embryo require an autonomously produ  
A.Reference number: A43903; MUID:92038439; PMID:1936564  
A.Accession: A43903  
A.Status: preliminary  
A.Molecule type: mRNA  
A.Residues: 'P', 633-1537, 'G', <WES>  
A.Cross-references: GB:564572; NID:g238616; PIDN:AAB20270.1, PID:g238617  
A.Note: sequence extracted from NCBI backbone (NCBIN:64572, NCBIP:64573)  
R.Venkatesan, M.; De Pablo, F.; Vogeli, G.; Simpson, R.T.  
Proc. Natl. Acad. Sci. U.S.A. 83, 3351-3355, 1986  
A.Title: Structure and developmentally regulated expression of a Strongylocentrotus purp  
A.Reference number: A23940; MUID:86205894; PMID:3458186  
A.Accession: A23940  
A.Molecule type: DNA  
A.Residues: 742-812 <VEN>  
A.Cross-references: EMBL:M13206  
C.Superfamily: collagen alpha 1(IV) chain  
C.Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix  
F:29-161/Domain: amino-terminal nonhelical, 7S <7S>  
F:162-1523/Region: interrupted helical  
F:1524-1752/Domain: carboxyl-terminal nonhelical, NC1 <NC1>  
F:1534-1634/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F:1644-1748/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F:129/Modified site: allysine (lys) #status predicted

Query Match	9.6%	Score 125.5;	DB 2;	Length 1752;
Best Local Similarity	25.0%	Pred. No. 0.0059;		
Matches 38;	Conservative 20;	Mismatches 51;	Indels 43;	Gaps 6;
QY	58	LQGPAGVGRDGSFGANGIPGTPTGIPGRDGFGRGKEGCELRFSF-----		100
Db	1471	LTGDKGTGPGVGNPCVSGVPEPGLKGRQGRGNQGGDGPDPOTKGEAGIPGSSSGFF		1530
QY	101	--EESWTFNRYKQC-----SW-----SLNYGIDGLKIAECFTPTKRSNSALRYL		142
Db	1531	ITRRSQTSTIPQCCPGGTAKMMHGVSLFVQGNERGHGQDLKPGSSC---LKRFTMPFL		1586
QY	143	FSGSLRLKCRNACCRCRWYFTFNGAECGSLPI		174
Db	1587	FC-NINNVCHVASRNDYSYMLSTTE---PMEH		1614
RESULT 4				
T15268				
hypochemical protein F59g12.12 - <i>Caenorhabditis elegans</i>				
Species: <i>Caenorhabditis elegans</i>				

C.Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C.Accession: F15268  
R.Johnson, D.  
submitted to the EMBL Data Library, May 1997  
A.Description: The sequence of C. elegans cosmid F59E12.  
A.Reference number: Z18318  
A.Accession: F15268  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: DNA  
A.Residues: 1-311 <JOH>  
A.Cross-references: UNIPROT:O01904; EMBL:AF003386; NID:G2088833; PID:G2088834; PIDN:AA85E12  
A.Experimental source: strain Bristol N2; clone F59E12  
C.Genetics:  
A.Gene: CESP:F59E12.12  
A.Map position: 2  
A.Introns: 24/2

## Query Match

Best Local Similarity	53.3%	Pred. No. 0.0013,
Matches 24; Conservative	4;	Mismatches 12; Indels 5; Gaps 1,

```

QY      55  GMC-----LQGPAGVPGRDGSPGANGIPGTPGIPGRDGFKEKGE      94
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      227  GKCDENVNAGP PPGSPGPPGLPGPDGLPCTPGNPGQDGEQGPAGE      271

```

## RESULT 5

collagen alpha 1(VII) chain - mouse (fragment)  
C|Species: Mus musculus (house mouse)  
C|Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C|Accession: A45748  
R|Li, K.; Christiano, A.M.; Copeland, N.G.; Gilbert, D.J.; Chu, M.L.; Jenkins, N.A.; Uitto  
Genomics 16, 733-739, 1993  
A|Title: cDNA cloning and chromosomal mapping of the mouse type VII collagen gene (col7a  
A|Reference number: A45748; MUID:93315168; PMID:8325648  
A|Accession: A45748  
A|Status: Preliminary  
A|Molecule type: nucleic acid  
A|Residues: 1-920 <L11>  
A|Cross-References: UNIPROT:Q63870; GB:S63654; NID:G386656; PIND:AMB27492.1; PID:G386657  
A|Experimental source: epidermal keratinocyte  
A|Note: sequence extracted from NCBI backbone (NCBIN:135000, NCBIPI:135001)

	Query Match	Score	DB	Length
Beet	Local Similarity	63.2%	Pred. No. 0.0054	
Match	24, Conservative	2, Mismatches	9, Indels	3, Gaps
OY	60	GPAGVPR---	DGSPGANGIPGTGPICPDRDFKKEKGE	94
DB	680	GPVGPVPGKAGDDGKPGKLNKNGKGDGDPEDDRKREKKGCD		717

## RESULT 6

type VII collagen - Chinese hamster (fragment)  
 C:Species: Cricetulus griseus (Chinese hamster)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 C:Accession: I48103  
 R:Greenspan, D.S.  
 Hum. Mol. Genet. 2, 273-278, 1993  
 A>Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous  
 A:Reference number: I48103; MUID:93271985; PMID:849916  
 A:Accession: I48103  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1549 <RES>  
 A:Cross-references: UNIPROT:Q60444; GB:I06863; NID:g388624; PIDN:AAA36968.1; PID:g3886255  
 F:1484-1536/Domain: animal kunitz-type proteinase inhibitor homology <BPI>

Query Match	9.48	Score	122.5	DB	2	Length	1549
Best Local Similarity	63.2%	Pred. No.	0.0096				
Matches	24	Conservative	2	Mismatches	9	Indels	3
						Gaps	1

OY 60 GPAGVPR---DSSPGANGIPGTPIRGDFGKGE 94  
|||||  
Db 426 GPGVPGKTGBDGKPGKNGKNGEPDGPEDGKGEKD 463

## RESULT 7

54657

collagen alpha 1(XIV) chain - human (fragments)

C/Species: Homo sapiens (man)

C/Date: 15-Jul-1995 #sequence revision 03-Nov-1995 #text\_change 09-Jul-2004

C/Accession: S4657; S37750; A44398

R/Brown, J.C.; Goldblk, R.; Mann, K.; Timpl, R.

Matrix Biol. 14, 287-295, 1994

A/Title: Structure and stability of the triple-helical domains of human collagen XIV.

A/Reference number: S4657; PMID:95128518; PMID:7827751

A/Accession: S4657

A/Status: preliminary

A/Molecule type: protein

A/Residues: 26-330 &lt;BRO&gt;

A/Cross-references: UNIPROT:Q9UDF6

R/Brown, J.C.; Goldblk, R.; Mann, K.; Timpl, R.

submitted to the EMBL Data Library, October 1993

A/Description: Structure and stability of the triple helical domains of human collagen X

A/Reference number: S37749

A/Accession: S37750

A/Molecule type: protein

A/Residues: 203-330 &lt;BR2&gt;

R/Brown, J.C.; Mann, K.; Wiedemann, H.; Timpl, R.

J. Cell Biol. 120, 557-567, 1993

A/Title: Structure and binding properties of collagen type XIV isolated from human plac

A/Reference number: A44398; PMID:93132085; PMID:8421066

A/Accession: A44398

A/Molecule type: protein

A/Residues: 1-17,18-25;26-42,'X',44-45;74-84;100-117;136-150,'SW',238-260,'X',262-263,'X'

A/Experimental source: placenta

C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit

ed and subsequently O-glycosylated.

C/Genetics:

A/Gene: COL14A1

C/Complex: type XIV collagen may be a heterotrimer of alpha 1(XIV) chains and one alpha

C/Function:

A/Description: structural component of extracellular fibrous polymer associated with col

C/Keywords: cold coil; extracellular matrix; glycoprotein; hydroxyproline; hydroxyprol

F;34,88,111,117,141,144,147,219,222,225,228,231,237,240,243,249,252,267,278,292,295,298,

F;3,206,289/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

F;43,261/Modified site: 5-hydroxylysine (Lys) #status experimental

F;43,261,264/Binding site: carbohydrate (Lys) (covalent) #status experimental

F;264/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental

Query Match 9.3%; Score 122; DB 2; Length 330;

Best Local Similarity 31.9%; Pred. No. 0.0019;

Matches 3; Conservative 10; Mismatches 33; Indels 34; Gaps 5;

OY 3 PGGRAASPOLLLGFLVLLLLQLSAPSSASENPKVKOKALIR--OREVVD-----L 52

Db 155 POGALGPPV-----PGAKGERGERDLSQAMVKKVAYQVEQLIQSMAR 201

OY 53 YNGN-----CLOGPAGVGRDGSFGA---NGIPGTPIPGRDGFKGE 94

Db 202 YNGIPSHSSIRTVQVGPGRGSGCGAGPQGPETPGPPGNAVGPTRPGE 254

## RESULT 8

A39762

collagen alpha 1(XIV) chain - bovine (fragments)

C/Species: Bos primigenius taurus (cattle)

C/Date: 14-Feb-1992 #sequence revision 05-Apr-1995 #text\_change 09-Jul-2004

C/Accession: A39762; S23815; S16905

R/Dubler, B.; van der Rest, M.

J. Biol. Chem. 266, 6853-6858, 1991

A/Title: Type XIV collagen, a new homotrimeric molecule extracted from fetal bovine skin

A/Reference number: A39762; PMID:91201330; PMID:2016301

A/Accession: A39762

A/Molecule type: protein

A/Residues: 1-19;20-61;62-71;72-81;82-123;124-178 &lt;DDB&gt;

A/Cross-references: UNIPROT:Q7M2W9

R/Aubert-Foucher, E.; Font, B.; Bichenberger, D.; Goldschmidt, D.; Leithae, C.; van der

J. Biol. Chem. 267, 15759-15764, 1992

A/Title: Purification and characterization of native type XIV collagen.

A/Reference number: S23815; PMID:92348439; PMID:1322405

A/Accession: S23815

A/Molecule type: protein

A/Residues: 'XR',82-94,'X',96 &lt;AUB&gt;

C/Complex: homotrimer

C/Keywords: cold coil, disulfide bond; extracellular matrix; homotrimer; hydroxylysine

Query Match 9.3%; Score 121; DB 2; Length 178;

Best Local Similarity 33.0%; Pred. No. 0.0012;

Matches 32; Conservative 7; Mismatches 34; Indels 24; Gaps 4;

OY 1 MHQGRASPOLLLGFLVLLLLQLSAPSSASENPKVKOKALIRREVVDLYNGMCTOG 60

Db 47 MGPQALGPP-----GVPEAGD---LDSQAVRQ-VCEQLIQSHITQG 85

OY 61 PAGVPRDGSFGA---NGIPGTPIPGRDGFKGE 94

Db 86 PPEPRPSPGTPRGSGPPGAGFPNGNVPSXGE 122

## RESULT 9

CGHUTL

collagen alpha 1(III) chain precursor - human

N/Alternate names: procollagen alpha 1(III) chain

C/Species: Homo sapiens (man)

C/Date: 24-Apr-1984 #sequence revision 01-Sep-1995 #text\_change 09-Jul-2004

C/Accession: S05272; S04642; PE0011; S01726; S04887; A90399; A94562; I51868; S59511; A9

R/Prockop, D.J.

submitted to the EMBL Data Library, February 1989

A/Reference number: S05272

A/Accession: S05272

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1240,'V',1242-1466 &lt;PRC&gt;

A/Cross-references: UNIPROT:P02461; EMBL:X14420; NID:g30057; PIDN:CAA32583.1; PID:g3005

R/Ala-Tokko, L.; Konluhaari, S.; Baldwin, C.T.; Kivaniemi, H.; Prockop, D.J.

Biochem. J. 260, 509-516, 1989

A/Title: Structure of cDNA clones coding for the entire prepro-alpha1(III) chain of hum

erences.

A/Reference number: S04642; PMID:89350838; PMID:2764886

A/Accession: S04642

A/Molecule type: mRNA

A/Residues: 1-1196 &lt;Ala&gt;

A/Cross-references: EMBL:X14420; NID:g30057; PIDN:CAA32583.1; PID:g30058

A/Note: the complete sequence is not shown

R/Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.

Gene 78, 255-265, 1989

A/Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene

A/Reference number: PE0011; PMID:89378752; PMID:2777083

A/Accession: PE0011

A/Molecule type: DNA

A/Residues: 1-176 &lt;BEN&gt;

A/Cross-references: GB:M26939; NID:g180813; PIDN:AAA22040.1; PID:g180814

R/Toman, P.D.; Ricca, G.A.; de Crombrughe, B.

Nucleic Acids Res. 16, 7201, 1988

A/Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human pr

A/Reference number: S01726; PMID:88303360; PMID:3405773

A/Accession: S01726

A/Molecule type: mRNA

A/Residues: 1-170 &lt;TOM&gt;

A/Cross-references: EMBL:X07240; NID:g30060; PIDN:CAA30229.1; PID:g30061

A/Note: the authors translated the codon CAG for residue 154 as His

R/Janezko, R.A.; Ramirez, F.

Nucleic Acids Res. 17, 6742, 1989

A/Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen

A/Reference number: S04887; PMID:89386015; PMID:2780304

A:Accession: S04887  
A:Molecule type: mRNA  
A:Residues: 149-163, 'G', 164-240, 'D', 242-471, 'D', 473-487, 'L', 489, 'S', 491-613, 'Y', 615-634,  
A:Cross-references: EMBL:X15332; NID:g29545; PIDN:CAA3387.1; PID:g930045  
A:Note: the authors' translation of residues 905-932 is inconsistent with the nucleotide  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 16, 1158-1164, 1977  
A:Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptide  
A:Reference number: A90399; MUID:77134724; PMID:557335  
A:Accession: A90399  
A:Molecule type: protein  
A:Residues: 'V', 169-225, 229-232, 'P', 234-292, 'D', 294-398 <SEY1>  
A:Experimental source: liver  
A:Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galact  
R:Sever, J.M.  
submitted to the Atlas, December 1977  
A:Reference number: A94562  
A:Accession: A94562  
A:Molecule type: protein  
A:Residues: 'V', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>  
A:Experimental source: liver  
A:Note: author submitted corrections to A90399  
R:Miliewicz, D.M.; Witz, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P.H.  
Am. J. Hum. Genet. 53, 62-70, 1993  
A:Title: Parental somatic and germ-line mosaicism for a multiection deletion with unusual  
splicing.  
A:Reference number: 151868; MUID:93304430; PMID:8317500  
A:Accession: 151868  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 186-194 <ML>  
A:Cross-references: GB:S62925; NID:g386425; PIDN:AAJ3937.1; PID:g4261637  
R:Chiodo, A.A.; Silience, D.O.; Cole, W.G.; Bateman, J.F.  
Biochem. J. 311, 939-943, 1995  
A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3  
A:Reference number: S59511; MUID:96067614; PMID:7487954  
A:Accession: S59511  
A:Molecule type: mRNA  
A:Residues: 102-423 <CHT>  
A:Cross-references: GB:S79877; NID:g1195576; PIDN:AAB5615.1; PID:g1195577  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 17, 3404-3411, 1978  
A:Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBr pe  
A:Reference number: A90414; MUID:79000343; PMID:687591  
A:Accession: A90414  
A:Molecule type: protein  
A:Residues: 399-675, 'N', 677-727 <SEY3>  
A:Experimental source: liver  
R:Lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.  
J. Biol. Chem. 266, 5256-5259, 1991  
A:Title: G to T transversion at position +5 of a splice donor site causes skipping of th  
A:Reference number: 155349; MUID:91161621; PMID:1672129  
A:Accession: 155349  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 537-605 <LEB>  
A:Cross-references: GB:M59312; NID:g180815; PIDN:AA52041.1; PID:g180816  
R:Sever, J.M.; Mainardi, C.; Kang, A.H.  
Biochemistry 19, 1588-1589, 1980  
A:Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB5 from cy  
A:Reference number: A90438; MUID:80198282; PMID:6246925  
A:Accession: A90438  
A:Molecule type: protein  
A:Residues: 728-895, 'A', 897-964 <SEY4>  
A:Experimental source: liver  
R:Cole, W.G.; Chiodo, A.A.; Lamanche, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Chan  
J. Biol. Chem. 265, 17070-17077, 1990  
A:Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping an  
A:Reference number: A38303; MUID:91009133; PMID:2145268  
A:Accession: A38303  
A:Molecule type: mRNA  
A:Residues: 861-1015 <COL>  
A:Cross-references: GB:J05617; GB:M55603; GB:M59227; NID:g180878; PIDN:AAB59383.1; PID:g

A:Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syn  
R:Mancoo, B.S.; Daigleish, R.  
Nucleic Acids Res. 16, 2337, 1988  
A:Title: Human pro alpha1(III) collagen: cDNA sequence for the 3' end.  
A:Reference number: S02119; MUID:88189827; PMID:3357782  
A:Accession: S02119  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>  
A:Cross-references: EMBL:X06700; NID:g30053; PIDN:CAA29886.1; PID:g30054  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 20, 2621-2627, 1981  
A:Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB9 from cy  
A:Reference number: A90446; MUID:81208139; PMID:7016180  
A:Accession: A90446  
A:Molecule type: protein  
A:Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 1157-  
A:Experimental source: liver  
R:Loidl, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.; Mye  
Nucleic Acids Res. 12, 9383-9394, 1984  
A:Title: Molecular cloning and carboxyl-propeptide analysis of human type III procollage  
A:Reference number: A93551; MUID:85087944; PMID:6096827  
A:Accession: A93551  
A:Molecule type: mRNA  
A:Residues: 1065-1155, 'P', 1157-1466 <LOI>  
A:Cross-references: EMBL:X01655; EMBL:X01742; NID:g29584; PIDN:CAA25821.1  
R:Meiklun, M.; Daigleish, R.; Kuve-Beckertman, B.; Remard, S.I.; Tolstoshev, P.; Brant  
Biochemistry 25, 1408-1413, 1986  
A:Title: Human type III collagen gene expression is coordinately modulated with the type  
A:Reference number: 152393; MUID:86187804; PMID:754462  
A:Accession: 152393  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1161-1200 <MS>  
A:Cross-references: GB:M13146; NID:g180415; PIDN:AA52003.1; PID:g180416  
R:Emanuel, B.S.; Cannizzaro, L.A.; Sever, J.M.; Myers, J.C.  
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985  
A:Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm  
A:Reference number: 159025; MUID:85157605; PMID:3858826  
A:Accession: 179359  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1165-1196 <EMA>  
A:Cross-references: GB:M1134; NID:g180417; PIDN:AA52004.1; PID:g180418  
R:Chu, M.L.; Weil, D.; de Wae, W.; Bernard, M.; Sipola, M.; Ramirez, F.  
J. Biol. Chem. 260, 4357-4363, 1985  
A:Title: Isolation of cDNA and genomic clones encoding human pro-alpha1(III) collagen. P  
A:Reference number: A92516; MUID:85157600; PMID:2579949  
A:Accession: A92516  
A:Molecule type: DNA  
A:Residues: 1176-1240, 'V', 1242-1356, 'P', 1358-1466 <CHU>  
A:Cross-references: GB:M10615; GB:M10793; GB:M10794; GB:M10795; GB:M10796; GB:M10797; GB  
A:Experimental source: liver  
A:Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given f  
action  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit (C  
3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently O  
C:Genes: GDB:COL3A1  
A:Gene: GDB:COL3A1  
A:Cross-references: GDB:118729; OMIM:120180  
A:Map position: 2q31-2q31  
A:Introns: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/3; 1275/1; 1337/3; 1418/3  
A:Note: the list of introns is incomplete; defects in this gene can result in Ehlers-Dan  
C:Complex: type III collagen is a homotrimer of monomers initially linked by disulfide b  
er of their length, is formed with desmosine cross-links made from lysine and allysine r  
C:Function:  
A:Description: structural component of extracellular fibrous polymer that maintains inte  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C:Keyword: coiled coil; Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hyd  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-153/Domain: amino-terminal propeptide #status predicted <PRO>  
F:1-91/Domain: von Willebrand factor type C repeat homology <WVC>  
F:154-1221/Product: collagen alpha 1(III) chain #status predicted <MAT>



F:154-167/Region: amino-terminal nonhelical telopeptide  
F:168-1196/Region: helical  
F:1091-1093/Region: cell attachment (R-G-D) motif  
F:1197-1221/Region: carboxyl-terminal nonhelical telopeptide  
F:1222-1466/Domain: fibrillar collagen propeptide #status predicted <CPR>  
F:1238-1466/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:153-154/Cleavage site: Pro-Gln (procollagen N-endopeptidase) #status predicted  
F:154/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:161-1212/Modified site: allysine (Lys) #status predicted  
F:263-284,860,977,1106/Modified site: 5-hydroxylysine (Lys) #status experimental  
F:263/Binding site: carboxyhydroxy (Lys) (covalent) #status experimental  
F:584-1099/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental  
F:948-949/Cleavage site: Gly-Ile (collagenase) #status experimental  
F:1106/Binding site: carboxyhydroxy (Lys) (covalent) #status predicted

Query Match 9.3%; Score 121; DB 1; Length 1466;  
Best Local Similarity 40.8%; Pred. No. 0.012;  
Matches 31; Conservative 7; Mismatches 30; Indels 8; Gaps 3;

QY 54 NMCICGPAVPRDGSFGANGIPGTGPIGRDGFKEKGECLRESF--ESGWTPIYKQ 110  
DB 100 NGGCPGPKDPPPPGIPGRNPDGIPGQSPSPSPGPGIC--ESCPGPGQVSPQYDS 157  
QY 111 CSWSINYGIDLKIA 126  
DB 158 YDYKS--GVAVGGLA 170

RESULT 10  
S16366  
collagen alpha 2(IV) chain precursor - pig roundworm  
C:Species: Ascaris suum (pig roundworm)  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: S16366  
R:Petit, J.; Kingston, I.B.  
J. Biol. Chem. 266, 16149-16156, 1991  
A:Title: The complete primary structure of a nematode alpha-2(IV) collagen and the parti  
A:Reference number: S16366; MUID:91340768; PMID:11714907  
A:Accession: S16366  
A:Molecule type: mRNA  
A:Residues: 1-1763 <JBI>  
A:Cross-references: UNIPROT:P27393; GB:M67507; NID:gl59648; PIDN:AAA18014.1; PID:gl59648  
C:Genetics: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;  
A:ntons: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfid  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-1763/Product: collagen alpha 2(IV) chain #status predicted <MNT>  
F:27-42/Domain: non-collagenous NH1 #status predicted <NH1>  
F:43-1529/Domain: collagenous #status predicted <COL>  
F:197-199/Region: cell attachment (R-G-D) motif  
F:1530-1763/Domain: carboxyl-terminal nonhelical, NCI #status predicted <NC1>  
F:1530-1638/Domain: repeat NCI #status predicted <NC11>  
F:1639-1763/Domain: repeat NCI #status predicted <NC12>  
F:31-34,39,41,536,539/Disulfide bonds: interchain #status predicted  
F:126/Binding site: carboxyhydroxy (Asn) (covalent) #status predicted  
F:1593-1599,1702-1709/Disulfide bonds: #status predicted

Query Match 9.3%; Score 121; DB 2; Length 1763;  
Best Local Similarity 62.9%; Pred. No. 0.015;  
Matches 22; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 60 GPAGVPRDGSFGANGIPGTGPIGRDGFKEKGE 94  
DB 162 GYAGAPGEKGDADGAPGPGIPGPRDGFKEKGD 196

RESULT 11  
S29328  
complement subcomponent C1q chain C - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S29328  
R:Petty, F.; Reid, K.B.M.; Loos, M.  
Eur. J. Biochem. 209, 129-134, 1992  
A:Title: Isolation, sequence analysis and characterization of cDNA clones coding for th  
e cecerebellin.  
A:Reference number: S29328; MUID:93011118; PMID:1396691  
A:Accession: S29328  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-246 <PET>  
A:Cross-references: UNIPROT:Q02105; EMBL:X66235; NID:G50228; PIDN:CAA46993.1; PID:G5022  
C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal ho  
F:122-245/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match 9.1%; Score 119.5; DB 2; Length 246;  
Best Local Similarity 36.6%; Pred. No. 0.0023;  
Matches 34; Conservative 6; Mismatches 32; Indels 21; Gaps 3;

QY 5 GRASPQLLGFVLVLLLLQLSAPSSASENPKVKALIRQEVVDLYNGMCLQGPAV 64  
DB 4 GRSCQFQ--CGICLLFLFLALPLRSQAS-----AGCTGIRGMPMPRA 45  
QY 65 PGRDGS--PGANGIPGTGPIGRDGFKEKGE 94  
DB 46 PGKDGHDLGQPKGEGRIPAVPQTGPKQKGE 78

RESULT 12  
T29731  
hypothetical protein F41F3.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T29731  
R:Badcock, K.; Le, T.T.; Gattung, S.  
Submitted to the EMBL Data Library, April 1996  
A:Description: The sequence of C. elegans cosmid F41F3.  
A:Reference number: Z20674  
A:Accession: T29731  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-310 <BAD>  
A:Cross-references: UNIPROT:Q20282; EMBL:U55366; PIDN:AAA97982.1; GSPDB:GN00023; CESP:F  
C:Experimental source: strain Bristol N2; clone F41F3  
C:Genetics: 27/3; 68/3  
A:Gene: CESP:F41F3.4  
A:Map position: 5  
A:ntons: 27/3; 68/3

Query Match 9.1%; Score 119.5; DB 2; Length 310;  
Best Local Similarity 40.7%; Pred. No. 0.003; 10; Indels 21; Gaps 1;  
Matches 24; Conservative 4; Mismatches 10; Indels 21; Gaps 1;

QY 57 CLOGPAGVPRDGSFGANGIPGTGPI-----PGRDGFKEKGE 94  
DB 158 CPGPAGPGEHGDGAPGAPGPGVGGSGSAPGAPGPKGAPGPGQPRDGFKEKGD 216

RESULT 13  
A61396  
collagen alpha 1(II) chain - golden hamster (Fragments)  
C:Species: Mesocricetus auratus (golden hamster)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Jul-2004  
C:Accession: A61396  
R:Cizdziel, P.E.; Hosi, J.; Montgomery, J.C.; Wiseman, R.W.; Barrett, J.C.  
Mol. Cell Biol. 4, 14-24, 1991  
A:Title: Loss of a tumor suppressor gene function is correlated with downregulation of  
A:Reference number: A61396; MUID:91182265; PMID:2009131  
A:Accession: A61396  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-323 <CTZ>  
A:Cross-references: UNIPROT:Q7M099  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology

F:120-323/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
Query Match 9.1%; Score 119.5; DB 2; Length 323;  
Best Local Similarity 20.6%; Pred. No. 0.0031;  
Matches 57; Conservative 20; Mismatches 47; Indels 153; Gaps 10;  
QY 58 LOGPAGVPRGDSRSGN-----GIPGPR----- 80  
DB 9 LTGPAGEPRGSGSGADGPPRGDGAAGVKDREGTALGAPGAPPPAGPAGATGKQG 68  
QY 81 -----GIPGRDGFKEGKE-----CLRESPE 101  
DB 69 DRGSGAGQPMGSPGAPGARGIPGPGRGDKSESGPSLRQHDVEVDATLKSUNQGE 128  
QY 102 ESMTP-----NYQC--SWSLANYGIDLGKIAECTTKRS----- 135  
DB 129 SIRSPGSGCKNPARTQDLKCHPEKWSGDYWDIPNQ--GCTLDARKAFQCMETGTCVY 186  
QY 136 -NSA-----LRLVLF---SGSLRLKCRNACQGRWYTFNGAGC 168  
DB 187 PNPASVPKKNWSSKKGEKKHIVQMTFLRLSTEGSONITTHCKN----- 231  
QY 169 SGPLPIEATVLDGSPELNSTINHTSSVEGLCEG 205  
DB 232 -----STAYLDEAAGNLKKALLIQSNDVEKRAEG 261  
RESULT 14  
CGHUB3  
collagen alpha 3(IV) chain precursor, long splice form - human  
N/Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form  
C/Species: Homo sapiens (man)  
C/Date: 28-Oct-1994 #sequence revision 03-Oct-1995 #text\_change 09-Jul-2004  
A/Accession: A54763; A43928; A44043; A45971; A39786  
R/Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tlygsvaason, K.; Reeders, S.T.  
J. Biol. Chem. 269, 23013-23017, 1994  
A/Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression  
A/Reference number: A54763; MUID:94364994; PMID:8083201  
A/Accession: A54763  
A/Molecule type: mRNA  
A/Residues: 1-1670 <MAR>  
A/Cross-references: UNIPROT:Q01955; GB:X80031; NID:g577563; PID:g577564  
A/Experimental source: kidney  
R/Furner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.  
J. Clin. Invest. 89, 597-601, 1992  
A/Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha  
A/Reference number: A43928; MUID:92147878; PMID:1737849  
A/Accession: A43928  
A/Molecule type: mRNA  
A/Residues: 1331-1524, 'I', 1526-1670 <TUR>  
A/Cross-references: GB:M81379  
A/Experimental source: kidney  
R/Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.  
J. Biol. Chem. 267, 19780-19784, 1992  
A/Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture  
A/Reference number: A44043; MUID:93015826; PMID:1400291  
A/Accession: A44043  
A/Molecule type: DNA, mRNA  
A/Residues: 1386-1670 <QUI>  
A/Cross-references: GB:M92993; NID:g177895; PIDN:AAA21610.1; PID:g177896  
A/Note: sequence extracted from NCBI backbone (NCBI:P115597)  
R/Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.  
J. Biol. Chem. 269, 17358, 1994  
A/Reference number: A44738; MUID:94274734; PMID:8006044  
A/Content: annotation; extratum; correction to intronic sequence in A44043  
R/Bernal, D.; Quinones, S.; Saus, J.  
J. Biol. Chem. 268, 12090-12094, 1993  
A/Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.  
A/Reference number: A45971; MUID:93280184; PMID:8505332  
A/Accession: A45971  
A/status: nucleic acid sequence not shown  
A/Molecule type: mRNA

A/Residues: 1427-1444 <BBR>  
A/Note: sequence extracted from NCBI backbone (NCBI:133363); sequence incorrectly ident  
R/Morrison, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reeders, S.T.  
Am. J. Hum. Genet. 49, 545-554, 1991  
A/Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of  
A/Reference number: A39786; MUID:91353570; PMID:1882840  
A/Accession: A39786  
A/Molecule type: mRNA  
A/Residues: 1453-1593, 'A', 1595-1670 <NOR>  
A/Cross-references: GB:S55790; NID:9234418; PIDN:AB19637.1; PID:9234419  
A/Comment: Prolines and lysines at the third position of the tripeptide repeating unit (ed and subsequently O-glycosylated.  
C/Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope  
C/Genetics:  
A/Genes: GDB:COL4A3  
A/Cross-references: GDB:128351; OMIM:120070  
A/Map position: 2q36-q37  
A/Intons: 1385/1; 1418/1; 1488/1; 1547/2; 1585/3; 1643/2 #status incomplete  
A/Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands with  
C/Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 3 (among trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a  
er associations in the interrupted helical domain (with disulfide and desmosine cross-11  
C/Function:  
A/Description: minor structural component of extracellular basement membrane in kidney  
C/Superfamily: collagen alpha 1(IV) chain  
C/Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracel  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MAR>  
F:29-42/Domain: amino-terminal nonhelical, NH1 <NH1>  
F:1438-1438/Region: interrupted helical  
F:791-793/Region: cell attachment (R-G-D) motif  
F:996-998/Region: cell attachment (R-G-D) motif  
F:1154-1156/Region: cell attachment (R-G-D) motif  
F:1306-1308/Region: cell attachment (R-G-D) motif  
F:1345-1347/Region: cell attachment (R-G-D) motif  
F:1432-1434/Region: cell attachment (R-G-D) motif  
F:1438-1670/Domain: carboxyl-terminal nonhelical, NCI <NC1>  
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F:31,33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status predi  
F:553/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:1466-1548, 1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted  
F:1508-1511, 1616-1622/Disulfide bonds: #status predicted  
F:1570-1662, 1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted  
Query Match 9.1%; Score 119.5; DB 1; Length 1670;  
Best Local Similarity 57.1%; Pred. No. 0.02;  
Matches 24; Conservative 3; Mismatches 12; Indels 3; Gaps 1;  
QY 59 OGPAGVPRGDSRSGNNGIPGTGIPGRD---GFKGEKCECLR 97  
DB 504 OGNAGIKGSPGSPGNTGLRPFPGAGQDPLKKEKERTLQ 545  
RESULT 15  
A54849  
collagen alpha 1(VII) chain precursor - human  
N/Alternate names: procollagen alpha 1(VII) chain  
C/Species: Homo sapiens (man)  
C/Date: 04-Nov-1994 #sequence revision 04-Nov-1994 #text\_change 09-Jul-2004  
A/Accession: A54849; PH0844; S16316; S16328; A30296; I84686  
R/Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.  
J. Biol. Chem. 269, 20256-20262, 1994  
A/Title: Cloning of human type VII collagen. Complete primary sequence of the alpha(VII  
A/Reference number: A54849; MUID:94327588; PMID:8051117  
A/Accession: A54849  
A/status: not compared with conceptual translation  
A/Molecule type: mRNA  
A/Cross-references: UNIPROT:Q02388; GB:I02870; NID:9987124; PIDN:AAA5438.1; PID:g987125  
R/Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.  
Biochem. Biophys. Res. Commun. 183, 958-963, 1992  
A/Title: Molecular cloning and characterization of type VII collagen cDNA.

A:Reference number: PH0844; MUID:92231902; PMID:1567409  
 A:Accession: PH0844  
 A:Molecule type: mRNA  
 A:Residues: 'EPR',340-475, 'RALSTASHSTICWDRATMHPNCRSHWTTRACEDCPNPASHRAARAG',524-528, 'C',  
 A:Cross-references: DDBJ:DJ1152; DDBJ:DJ1694; NID:9455698; PIDN:BA02853.1; PID:9453699  
 A:Experimental source: Keratinocyte  
 A:Note: the authors translated the codon ACC for residues 394 and 397 as Tyr  
 R:Parente, M.G.; Chung, L.C.; Rymaszewski, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mat  
 Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991  
 A:Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.  
 A:Reference number: S16316; MUID:91334380; PMID:1871109  
 A:Accession: S16316  
 A:Molecule type: mRNA  
 A:Residues: 815-892, 'E',894-1439, 'PAR'-  
 A:Cross-references: GB:M65158; GB:S49017; NID:q180914; PIDN:AAA96439.1; PID:q180915  
 A:Experimental source: Keratinocyte  
 R:Gammone, W.R.; Abernethy, M.L.; Padilla, K.M.; Pridayan, P.S.; Cook, M.E.; Wright, J.;  
 J. Invest. Dermatol. 99, 691-696, 1992  
 A:Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot  
 A:Reference number: 156328; MUID:93107742; PMID:1469284  
 A:Accession: 156328  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 'EPR',372-517, 'DV',520-540, 'W',542-1255, 'RES'-  
 A:Cross-references: GB:S51336; NID:q262308; PIDN:AB24637.1; PID:q262309  
 R:Seitler, J.L.; Eissen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgess, R.E.  
 J. Biol. Chem. 264, 3822-3826, 1989  
 A:Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagenase  
 A:Reference number: A30296; MUID:9139437; PMID:2537292  
 A:Accession: A30296  
 A:Molecule type: protein  
 A:Residues: 'A',1240-1246, 'G',1248-1250, 'XE',1253-1255, 'Q',1257, 'E',2032, 'C',2034-2041, '  
 A:Note: two reported peptides cannot be reliably located  
 R:Greenspan, D.S.  
 Hum. Mol. Genet. 2, 273-278, 1993  
 A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous  
 A:Reference number: 148103; MUID:93271985; PMID:8499916  
 A:Accession: 184686  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 2395-2871, 'S',2873-2944, 'RE2'-  
 A:Cross-references: GB:L06862; NID:q388713; PIDN:AAA89196.1; PID:q388714  
 R:Christiano, A.M.; Rymaszewski, M.; Uitto, J.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994  
 A:Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs  
 A:Reference number: A55255; MUID:94224777; PMID:8170945  
 A:Accession: A55255  
 A:Contents: annotation  
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 ed and subsequently O-glycosylated.  
 C:Genetics:  
 A:Gene: GDB:COL7A1; EBR1; EBD1; EB  
 A:Cross-references: GDB:128750; OMIM:120120  
 A:Map position: 3p21.3-3p21.3  
 A:Note: defects in this gene can result in dominant and recessive dystrophic epidermolys  
 A:Note: there are 118 introns  
 C:Complex: type VII collagen is probably a homotrimer  
 C:Function:  
 A:Description: Structural component of extracellular polymer associated with anchoring  
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>  
 F:17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>  
 F:36-201/Domain: von Willebrand factor type A repeat homology <VWA1>  
 F:327-413/Domain: fibronectin type III repeat homology <FN1>  
 F:414-502/Domain: fibronectin type III repeat homology <FN2>  
 F:508-593/Domain: fibronectin type III repeat homology <FN3>  
 F:598-683/Domain: fibronectin type III repeat homology <FN4>  
 F:686-771/Domain: fibronectin type III repeat homology <FN5>  
 F:776-862/Domain: fibronectin type III repeat homology <FN6>  
 F:864-952/Domain: fibronectin type III repeat homology <FN7>  
 F:954-1045/Domain: fibronectin type III repeat homology <FN8>  
 F:1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>

F:1170-1172/Region: cell attachment (R-G-D) motif  
 F:1189-1253/Region: cysteine/proline-rich  
 F:1254-2783/Region: interrupted helical  
 F:1334-1336/Region: cell attachment (R-G-D) motif  
 F:2008-2010/Region: cell attachment (R-G-D) motif  
 F:2553-2555/Region: cell attachment (R-G-D) motif  
 F:2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>  
 F:2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
 F:337,786,1109/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status expe  
 F:2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental  
 F:2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental  
 F:2634,2802,2804/Disulfide bonds: interchain #status predicted

Query Match 9.1%; Score 119.5; DB 2; Length 2944;  
 Best Local Similarity 60.5%; Pred. No. 0.037;  
 Matches 23; Conservative 3; Mismatches 9; Indels 3; Gaps 1;  
 QY 60 GPAGVPCR---DQSPGANGIPGTPGIPGRDGRKGEKE 94  
 DB 1824 GPEGLPGKPEDGKPELNGKNBPGDGEDGRKGEKD 1861

Search completed: August 4, 2005, 04:22:28  
 Job time : 20.5359 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:04:29 ; Search time 94.4771 Seconds  
(without alignments)  
1002.955 Million cell updates/sec

Title: US-10-634-108-2

Perfect score: 1307

Sequence: 1 MHPQGRAPQLLGLFLV.....GDASTGMSVSRRIIELEPK 245

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

A\_Geneseq\_16dc04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1307	100.0	245	5	ABBB0977 Rat REMOD
2	1307	100.0	245	6	ABG75757 Rat REMOD
3	1307	100.0	277	5	ABBB0979 Long form
4	1307	100.0	277	6	ABG75759 Rat REMOD
5	1287.5	98.5	244	8	ADO28703 HA4, SEQ
6	1225	93.7	243	4	AAB36667 Human sec
7	1225	93.7	243	5	ABG96340 Human ova
8	1225	93.7	243	5	ABJ05554 Breast ca
9	1225	93.7	243	5	ABBB0978 Human REM
10	1225	93.7	243	6	ABR85856 Human can
11	1225	93.7	243	6	ABR48227 Human b1a
12	1225	93.7	243	6	ABG75758 Human REM
13	1225	93.7	243	6	ABU56607 Lung canc
14	1225	93.7	243	7	ADBB0510 Ovarian c
15	1225	93.7	243	7	ADN39855 Cancer/an
16	1225	93.7	243	7	ADN38732 Cancer/an
17	1225	93.7	243	8	ADL70254 LBP1301 p
18	1225	93.7	278	3	AAB08856 Amino aci
19	1225	93.7	278	4	AAM5746 Human pro
20	1225	93.7	278	4	ABG96338 Human ova
21	1225	93.7	278	6	ABR47627 Breast ca
22	1222	93.5	243	3	AAJ91529 Human sec
23	1222	93.5	243	8	ADL71601 Novel hum
24	1222	93.5	278	5	ABP68631 Human pan
25	1220	93.3	243	6	ABJ37031 Human bre

#### ALIGNMENTS

26	1216	93.0	243	3	AAJ99462 Human PRO
27	1216	93.0	243	3	AAJ99441 Human PRO
28	1216	93.0	243	4	AAJ66190 Protein O
29	1216	93.0	243	4	AAJ66211 Protein O
30	1216	93.0	243	4	AAJ29206 Human PRO
31	1216	93.0	243	4	AAJ87586 Human PRO
32	1216	93.0	243	5	ABG95911 Human sec
33	1216	93.0	243	5	ABBB4939 Human PRO
34	1216	93.0	243	5	AAJ20462 Human tum
35	1216	93.0	243	5	ABJ55545 Human ang
36	1216	93.0	243	5	ABP68636 Human pan
37	1216	93.0	243	5	ABG78938 Human bre
38	1216	93.0	243	6	ABU58582 Human PRO
39	1216	93.0	243	6	ABU88130 Novel hum
40	1216	93.0	243	6	ABU84445 Human sec
41	1216	93.0	243	6	ABR66319 Human sec
42	1216	93.0	243	6	ABR65709 Human sec
43	1216	93.0	243	6	ABU96649 Human sec
44	1216	93.0	243	6	ABJ37761 Human tum
45	1216	93.0	243	6	ABU82888 Human PRO

RESULT 1  
ID ABB80977 standard; protein, 245 AA.  
AC ABB80977;  
DT 21-OCT-2002 (first entry)  
XX  
DE Rat REMODELIN polypeptide.  
XX  
KW REMODELIN; vulnery; vasotropic; cytostatic; osteopathic; collagen;  
XX gene therapy; bone; rat.  
XX  
OS Rattus sp.  
XX  
PN WO200242487-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 19-OCT-2001; 2001WO-US050940.  
XX  
PR 19-OCT-2000; 2000US-00692081.  
XX  
PA (MAIN-) MAINE MEDICAL CENT RES INST.  
XX  
PI Linder V, Friesel R;  
XX  
DR WPI; 2002-590472/63.  
XX  
NR N-PSDB; ABB86480.  
XX  
PT New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
XX negative remodelling, fibrosis, collagen disease, and bone disease in a  
XX mammal.  
XX  
PS Claim 5; Fig 4B; 187BP; English.  
XX

The invention relates to mammalian REMODELIN polypeptides and encoding polynucleotides. REMODELIN is used to diagnose arterial restenosis, negative remodelling or fibrosis, bone disease such as osteogenesis imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are used to affect cellular gene expression in a mammal, where the cellular gene is from transforming growth factor (TGF)-beta1, collagen fibrilaphal, osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein 4. The expression of osteopontin is dependent on Cbfa1. Compositions comprising antisense REMODELIN sequences are useful for treating diseases mediated by abnormal expression of a REMODELIN molecule in a human such as impaired wound healing, fibrosis of an organ, ectopic ossification, or

CC hypertrophic scar formation. REMODELIN is useful in the development of  
 CC bone during mammalian embryogenesis, proliferation and/or migration,  
 CC cellular signaling, adventitial fibrosis, negative remodeling and arterial  
 CC restenosis, smooth muscle cell proliferation, and arterial remodeling and  
 CC useful in gene therapy. REMODELIN is useful for elucidating the function  
 CC of REMODELIN molecules in a cell, to identify a compound that affects  
 CC REMODELIN expression and/or TGF-beta signaling, as a potential  
 CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
 CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
 CC formation, to promote bone fracture healing, and to increase or decrease  
 CC expression of REMODELIN in mammal. It is also useful for producing  
 CC recombinant cell and transgenic non-human mammals which are useful tools  
 CC for the study of REMODELIN action, for identifying novel diagnostic and  
 CC therapeutics for treatment, and for elucidating the cellular roles of  
 CC REMODELIN. The present sequence represents a rat REMODELIN polypeptide  
 XX  
 XX

SQ Sequence 245 AA;  
 Query Match 100.0%; Score 1307; DB 5; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3,7e-123;  
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPOGRAASPQLLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60  
 DB 1 MHPOGRAASPQLLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60  
 QY 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGECLEARSFESWTPNYKQCSWSSLYNGI 120  
 DB 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGECLEARSFESWTPNYKQCSWSSLYNGI 120  
 QY 121 DGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTFNAGCSGPIPIEAIITL 180  
 DB 121 DGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTFNAGCSGPIPIEAIITL 180  
 QY 181 DOGSEPLNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITL 240  
 DB 181 DOGSEPLNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITL 240  
 QY 241 EELPK 245  
 DB 241 EELPK 245

RESULT 2  
 ABG75757  
 ID ABG75757 standard; protein; 245 AA.

AC ABG75757;  
 XX  
 DT 28-APR-2003 (first entry)  
 XX  
 DE Rat REMODELIN protein.  
 XX

KM Rat; adventitia-inducible; REMODELIN; REMODEL;  
 KM adventitia induced bone expressed molecule; AIBE; TGF-beta;  
 KM transforming growth factor beta; adventitia; vascular remodeling;  
 KM restenosis; vascular injury; antisense therapy; TGF-beta signaling;  
 KM TGF-beta receptor type II; arterial remodeling; bone formation;  
 KM cartilage formation; osteogenesis imperfecta; Bethlem myopathy;  
 KM dystrophic epidermolysis bullosa; negative remodeling; wound healing;  
 KM arterial stenosis; fibrosis; calcification; transplant;  
 KM heart valve transplant; osteopathic; antiteriosclerotic; vulnerary.

OS Rattus sp.  
 XX  
 XX US2002161211-A1.  
 PN  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 19-OCT-2001; 2001US-00045992.  
 XX  
 XX 19-OCT-2000; 2000US-0062081.  
 PR  
 XX

PA (LIND/) LINDNER V.  
 PA (FRIE/) FRIESEL R.E.  
 XX  
 XX Lindner V, Friesel RE;  
 PI  
 XX  
 DR MPI: 2003-238238/23.  
 DR N-PSDB; ABX11340.  
 XX  
 XX  
 PT New isolated REMODELIN nucleic acid and polypeptide, useful for mediating  
 PT arterial remodeling, formation of bone and cartilage, and the diagnosis  
 PT and treatment of disorders associated with aberrant expression of  
 PT REMODELIN.  
 XX  
 PS Claim 3; Fig 4B; 81pp; English.

XX  
 XX The invention discloses an isolated nucleic acid encoding a mammalian  
 CC adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL  
 CC and/or adventitia induced bone expressed molecule, AIBE). REMODELIN  
 CC expression was induced by transforming growth factor beta (TGF-beta),  
 CC which is important because proliferative events occurring in the  
 CC adventitia contribute to vascular remodeling and restenosis in response  
 CC to vascular injury and TGF-beta has been shown to be a factor involved in  
 CC this. Also disclosed is an antibody raised against REMODELIN and methods  
 CC for treating a disease mediated by abnormal expression of a REMODELIN in  
 CC a human, for identifying a compound that affects or reduces expression of  
 CC REMODELIN in a cell (e.g. antisense therapy), for identifying a compound  
 CC that affects TGF-beta signaling and for increasing or reducing REMODELIN  
 CC expression in a mammal, comprising administering a REMODELIN expression  
 CC increasing or reducing TGF-beta to the mammal, thereby increasing  
 CC REMODELIN expression or inhibiting signaling via the TGF-beta receptor  
 CC type II and reducing expression of REMODELIN in the mammal. The methods  
 CC and compositions of the present invention are useful for mediating  
 CC arterial remodeling, formation of bone and cartilage and the diagnosis  
 CC and treatment of disorders associated with aberrant expression of  
 CC REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis  
 CC bullosa, Bethlem myopathy, negative remodeling, wound healing, arterial  
 CC stenosis, vessel injury, fibrosis and calcification of a transplant,  
 CC preferably a heart valve transplant. The sequence presented is the rat  
 CC REMODELIN protein  
 XX  
 XX

SQ Sequence 245 AA;  
 Query Match 100.0%; Score 1307; DB 6; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 3,7e-123;  
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPOGRAASPQLLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60  
 DB 1 MHPOGRAASPQLLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60  
 QY 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGECLEARSFESWTPNYKQCSWSSLYNGI 120  
 DB 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGECLEARSFESWTPNYKQCSWSSLYNGI 120  
 QY 121 DGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTFNAGCSGPIPIEAIITL 180  
 DB 121 DGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTFNAGCSGPIPIEAIITL 180  
 QY 181 DOGSEPLNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITL 240  
 DB 181 DOGSEPLNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITL 240  
 QY 241 EELPK 245  
 DB 241 EELPK 245

RESULT 3  
 ABB80979  
 ID ABB80979 standard; protein; 277 AA.  
 XX  
 AC ABB80979;  
 XX

DT 21-OCT-2002 (first entry)  
 XX Long form of rat REMODELIN polypeptide.  
 XX REMODELIN; vulnerability; vasotropic; cytostatic; osteopathic; collagen;  
 KM gene therapy; bone; rat.  
 XX  
 OS Rattus sp.  
 PN WO200242487-A2.  
 XX  
 PD 30-MAY-2002.  
 XX  
 PF 19-OCT-2001; 2001WO-US050940.  
 XX  
 PR 19-OCT-2000; 2000US-00692081.  
 XX  
 PA (MAIN-) MAINE MEDICAL CENT RES INST.  
 XX  
 PI Linder V, Friesel R;  
 DR MPI; 2002-590472/63.  
 DR N-PSDB; ABN86480.  
 XX  
 PT New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
 PT negative remodeling, fibrosis, collagen disease, and bone disease in a  
 mammal.  
 XX  
 PS Claim 5; Fig 4C; 187pp; English.  
 XX  
 CC The invention relates to mammalian REMODELIN polypeptides and encoding  
 CC polynucleotides. REMODELIN is used to diagnose arterial restenosis,  
 CC negative remodeling or fibrosis, bone disease such as osteogenesis  
 CC imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis  
 CC bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are  
 CC used to affect cellular gene expression in a mammal, where the cellular  
 CC gene is from transforming growth factor (TGF)-beta1, collagen Itriplehel,  
 CC osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein  
 CC 4. The expression of osteopontin is dependent on Cbfa1. Compositions  
 CC comprising antisense REMODELIN sequences are useful for treating diseases  
 CC mediated by abnormal expression of a REMODELIN molecule in a human such  
 CC as impaired wound healing, fibrosis of an organ, ectopic ossification, or  
 CC hypertrophic scar formation. REMODELIN is useful in the development of  
 CC bone during mammalian embryogenesis, proliferation and/or migration,  
 CC cellular signaling, adventitial fibrosis, negative remodeling and arterial  
 CC restenosis, smooth muscle cell proliferation, and arterial remodeling and  
 CC useful in gene therapy. REMODELIN is useful for elucidating the function  
 CC of REMODELIN molecules in a cell, to identify a compound that affects  
 CC REMODELIN expression and/or TGF-beta signaling, as a potential  
 CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
 CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
 CC formation, to promote bone fracture healing, and to increase or decrease  
 CC expression of REMODELIN in mammal. It is also useful for producing  
 CC recombinant cell and transgenic non-human mammals which are useful tools  
 CC for the study of REMODELIN action, for identifying novel diagnostics and  
 CC therapeutics for treatment, and for elucidating the cellular roles of  
 CC REMODELIN. The present sequence represents the long form of rat REMODELIN  
 CC polypeptide  
 CC  
 SQ Sequence 277 AA;  
 XX  
 XX  
 Query Match 100.0%; Score 1307; DB 5; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 4, 4e-123;  
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MHFQGRASPOLLIIGFLVILLILQSSASSASSENPKVKOKALIRORVVLDYNGKCLQG 60  
 DB 33 MHFQGRASPOLLIIGFLVILLILQSSASSASSENPKVKOKALIRORVVLDYNGKCLQG 92  
 QY 61 PACVPRDQSPGANGIPGTPGIRGDPFKGEKCECLRESFEESWTNYYKQCSWSSLNYGI 120  
 DB 93 PACVPRDQSPGANGIPGTPGIRGDPFKGEKCECLRESFEESWTNYYKQCSWSSLNYGI 152

QY 121 DLGKIAECTFTKRSNSAARVLPFSGSLRKLKGNACCQRRVFTFNGACSGPLPIEATITVL 180  
 DB 153 DLGKIAECTFTKRSNSAARVLPFSGSLRKLKGNACCQRRVFTFNGACSGPLPIEATITVL 212  
 QY 181 DQSPFLNSTINIHRTSVVEGLCEGIGAGLVVAIVWGCSPYPKKDASTGMNSVRRIT 240  
 DB 213 DQSPFLNSTINIHRTSVVEGLCEGIGAGLVVAIVWGCSDYPKKDASTGMNSVRRIT 272  
 QY 241 EELPK 245  
 DB 273 EELPK 277  
 RESULT 4  
 ID ABG75759 standard; protein; 277 AA.  
 XX  
 AC ABG75759;  
 XX  
 DT 28-APR-2003 (first entry)  
 XX  
 DE Rat REMODELIN long form (rREMODELINL) protein.  
 XX  
 KM Rat; adventitia-inducible; REMODELIN; rREMODELINL; REMODEL;  
 KM transformati growth factor beta; adventitia; vascular remodeling;  
 KM restenosis; vascular injury; antisense therapy; TGF-beta signaling;  
 KM TGF-beta receptor type II; arterial remodeling; bone formation;  
 KM cartilage formation; osteogenesis imperfecta; Bethlem myopathy;  
 KM dystrophic epidermolysis bullosa; negative remodeling; wound healing;  
 KM arterial stenosis; fibrosis; calcification; transplant;  
 KM heart valve transplant; osteopathic; antihypertensive; vulnerability.  
 XX  
 OS Rattus sp.  
 XX  
 PN US2002161211-A1.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 19-OCT-2001; 2001US-00045992.  
 XX  
 PR 19-OCT-2000; 2000US-00692081.  
 XX  
 PA (LIND/) LINDER V.  
 PA (FRIE/) FRIESEL R E.  
 PI Linder V, Friesel R E;  
 DR MPI; 2003-238238/23.  
 DR N-PSDB; ABX11340.  
 XX  
 PT New isolated REMODELIN nucleic acid and polypeptide, useful for mediating  
 PT arterial remodeling, formation of bone and cartilage, and the diagnosis  
 PT and treatment of disorders associated with aberrant expression of  
 PT REMODELIN.  
 XX  
 PS Claim 3; Fig 4C; 81pp; English.  
 XX  
 CC The invention discloses an isolated nucleic acid encoding a mammalian  
 CC adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL  
 CC and/or adventitia induced bone expressed molecule, AIBE). REMODELIN  
 CC expression was induced by transforming growth factor beta (TGF-beta),  
 CC which is important because proliferative events occurring in the  
 CC adventitia contribute to vascular remodeling and restenosis in response  
 CC to vascular injury and TGF-beta has been shown to be a factor involved in  
 CC this. Also disclosed is an antibody raised against REMODELIN and methods  
 CC for treating a disease mediated by abnormal expression of a REMODELIN in  
 CC a human, for identifying a compound that affects or reduces expression of  
 CC REMODELIN in a cell (e.g. antisense therapy), for identifying a compound  
 CC that affects TGF-beta signaling and for increasing or reducing REMODELIN  
 CC expression in a mammal, comprising administering a REMODELIN expression  
 CC increasing or reducing TGF-beta to the mammal, thereby increasing  
 CC REMODELIN expression or inhibiting signaling via the TGF-beta receptor

CC type II and reducing expression of REMODELIN in the mammal. The methods  
 CC and compositions of the present invention are useful for mediating  
 CC arterial remodeling, formation of bone and cartilage and the diagnosis  
 CC and treatment of disorders associated with aberrant expression of  
 CC REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis  
 CC bullosae, Bethlem myopathy, negative remodeling, wound healing, arterial  
 CC stenosis, vessel injury, fibrosis and calcification of a transplant,  
 CC preferably a heart valve transplant. The sequence presented is the long  
 CC form of the rat REMODELIN protein (rREMODELIN).

XX Sequence 277 AA;

Query Match 100.0%; Score 1307; DB 6; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-123; Indels 0; Gaps 0;  
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQRAASPOLLGLFLVLLLLQLSAPSSASENPKVKQALIRQREVVDLYNGMCLQG 60  
 DB 33 MHPQRAASPOLLGLFLVLLLLQLSAPSSASENPKVKQALIRQREVVDLYNGMCLQG 92  
 QY 61 PAGVPRDQSPGANGIPGTPGIPGRDGFGEKGECELRSEFESWTNPKQCSWSSLYNGI 120  
 DB 93 PAGVPRDQSPGANGIPGTPGIPGRDGFGEKGECELRSEFESWTNPKQCSWSSLYNGI 152  
 QY 121 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQWYFTFNAGCSGPIPIEAIYYL 180  
 DB 153 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQWYFTFNAGCSGPIPIEAIYYL 212  
 QY 181 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNSVSR111 240  
 DB 213 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNSVSR111 272  
 QY 241 EELPK 245  
 DB 273 EELPK 277

RESULT 5

ADO28703 standard; protein; 244 AA.

AD028703;  
 12-AUG-2004 (first entry)  
 HA4, SEQ ID 2.  
 Osteopathic; Cytostatic; Antiinflammatory; Antiarthritic; Antirheumatic;  
 bone degenerative disease; osteoporosis;  
 glucocorticoid induced osteoporosis; Paget's disease;  
 periodontal disease; tooth loss; bone fracture; rheumatoid arthritis;  
 periosteal osteolysis; osteogenesis imperfecta;  
 metastatic bone disease; hypercalcemia of malignancy.  
 Unidentified.  
 MO2004041205-AA2.  
 21-MAY-2004.  
 04-NOV-2003; 2003WO-US035139.  
 04-NOV-2002; 2002US-0423690P.  
 (TEXA ) UNIV TEXAS SYSTEM.  
 De Crombrughe B, Akiyama H;  
 WPI; 2004-411407/38.  
 N-PSDB; ADO28702.  
 Novel isolated osteoblast- and chondrocyte-specific small secreted

PT peptide referred as HA4 polypeptide, useful for diagnosing or treating  
 PT bone degenerative disease, and for screening drugs that alter bone  
 PT formation.

XX Claim 1; SEQ ID NO 2; 103bp; English.

CC The present invention relates to an isolated osteoblast- and chondrocyte-  
 CC specific small secreted peptide named HA4 polypeptide (I; ADO28703) and  
 CC its coding sequence (II; ADO28702). (I) or (II) are useful for diagnosing  
 CC or treating a bone degenerative disease in a subject such as  
 CC osteoporosis, glucocorticoid induced osteoporosis, Paget's disease,  
 CC abnormally increased bone turnover, periodontal disease, tooth loss, bone  
 CC fractures, rheumatoid arthritis, periosteal osteolysis, osteogenesis  
 CC imperfecta, metastatic bone disease and hypercalcemia of malignancy. (II)  
 CC is also useful for identifying a subject at risk of or suffering from a  
 CC bone degenerative disease.

XX Sequence 244 AA;

Query Match 98.5%; Score 1287.5; DB 8; Length 244;  
 Best Local Similarity 99.2%; Pred. No. 3,4e-121; Indels 1; Gaps 1;  
 Matches 243; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 MHPQRAASPOLLGLFLVLLLLQLSAPSSASENPKVKQALIRQREVVDLYNGMCLQG 60  
 DB 1 MHPQRAASPOLLGLFLVLLLLQLSAPSSASENPKVKQALIRQREVVDLYNGMCLQG 59  
 QY 61 PAGVPRDQSPGANGIPGTPGIPGRDGFGEKGECELRSEFESWTNPKQCSWSSLYNGI 120  
 DB 60 PAGVPRDQSPGANGIPGTPGIPGRDGFGEKGECELRSEFESWTNPKQCSWSSLYNGI 119  
 QY 121 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQWYFTFNAGCSGPIPIEAIYYL 180  
 DB 120 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQWYFTFNAGCSGPIPIEAIYYL 179  
 QY 181 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNSVSR111 240  
 DB 180 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNSVSR111 239  
 QY 241 EELPK 245  
 DB 240 EELPK 244

RESULT 6

AAB36667 standard; protein; 243 AA.

AAB36667;  
 14-MAR-2001 (first entry)  
 Human secretory protein TGC-628 SEQ ID NO:7.  
 Human; secretory protein; cancer; immune disease; infectious disease;  
 lung function disorder; liver function disorder; antiinflammatory;  
 gastrointestinal disorder; cytostatic; haematopoietic; anticoagulant;  
 immunomodulatory; hepatotropic; cell proliferation-stimulant;  
 cell migratory agent; cell differentiation-inducer.  
 Homo sapiens.  
 MO200071581-A1.  
 30-NOV-2000.  
 19-MAY-2000; 2000WO-JP003221.  
 20-MAY-1999; 99JP-00140229.  
 (TAKE ) TAKEDA CHEM IND LTD.  
 Itoh Y, Mogi S, Tanaka H, Ohkubo S, Ogi K;



```
XX WPI: 2001-032023/04.
DR N-PSDB; AAC90707.
PT Novel secretory protein and its salt with e.g. anti-cancer, anti-
XX inflammatory and hematopoietic, effects, applicable as drugs in remedies
XX and preventives to treat diseases like cancer and immune diseases.
PS Claim 1; Page 89-90; 122pp; Japanese.
CC AAC90701 to AAC90715 encode the human secretory proteins given in
CC AAB3661 to AAB3675. The proteins can have cytostatic, anti-
CC inflammatory, haematopoietic, anti-coagulant, immunomodulatory and
CC hepatotropic activities, and can be used as cell migratory agents, cell
CC proliferation-stimulants and cell differentiation-inducers. The proteins
CC are useful in the treatment and prevention of diseases such as cancer,
CC lung function disorder, liver function disorder, gastrointestinal
CC disorder and immune diseases. AAC90716 to AAC90755 represent PCR primers
CC which are used in the exemplification of the present invention
XX
SQ Sequence 243 AA;
Query Match 93.7%; Score 1225; DB 4; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.9e-115;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
QY 1 MHPOGRASPOLLLGLFLVLLLLQLSAPSSASBNPKYKOKALRQREVVLDYNGMCLQG 60
DB 1 MRPOGRASPOLRRG--LTLTLLTLLQLPAPSSASBIPKQKOKALRQREVVLDYNGMCLQG 58
QY 61 PAGVGRDSSPGANGIPGTGIPGRDGFKEGKCECLRSFESWTPNYKQCSWSSLYNGI 120
DB 59 PAEVPGRDSSPGANGIPGTGIPGRDGFKEGKCECLRSFESWTPNYKQCSWSSLYNGI 118
QY 121 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTNGAECGSPLEIAIYVL 180
DB 119 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTNGAECGSPLEIAIYVL 178
QY 181 DOGSPELNSTINIHRTSSVSGLCBEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRIT 240
DB 179 DOGSPENNSTINIHRTSSVSGLCBEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRIT 238
QY 241 EELPK 245
DB 239 EELPK 243
RESULT 7
ABG96340
ID ABG96340 standard; protein; 243 AA.
XX
AC ABG96340;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker M450.
XX
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nondercutaneous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker.
XX
OS Homo sapiens.
XX
PN MO200271928-A2.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002MO-US007826.
XX
```

```
PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324967P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX
XX (MILL-) MILLENNium PHARM INC.
XX
XX Monahan JE, Ganavarapu M, Hoersch S, Kamatkar S, Kovatis SG,
XX Meyers RE, Morrissey MP, Qandt PJ, Sen A, Vieby PO, Mills GB,
XX Bast RC, Lu K, Schmandt RE, Zhao X, Glatz K;
DR WPI: 2002-723277/78.
DR N-PSDB; AB576436.
XX
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
XX assessing the stage or progression of the disease, comprises comparing
XX the expression level of a cancer marker in a sample from a patient and
XX from a non cancer patient.
XX
XX Disclosure; Page 252; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
XX patient is afflicted with ovarian cancer. The method involves comparing
XX the expression level of a marker in a patient sample and the normal level
XX of expression of the marker in a control non-ovarian cancer sample, where
XX the marker is selected from 363 cancer markers described in the
XX specification. The method of the invention is useful in diagnosing or
XX characterizing cancer, in detecting the presence of cancer as early as
XX possible, and the recurrence of ovarian cancer. The method may also be of
XX particular use with patients having an enhanced risk of developing
XX ovarian cancer (e.g. patients having a familial history of ovarian
XX cancer). The cancer markers may be used in the management and treatment
XX of e.g. brain and central nervous system disorders (e.g. bacterial and
XX viral meningitis, Alzheimer's disease or Parkinson's disease), brain
XX disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
XX inflammations (e.g. bacterial or viral meningitis or encephalitis),
XX testicular disorders (e.g. nondercutaneous granulomatous orchitis),
XX connective tissue disorders, or heart disorders (e.g. ischaemic heart
XX disease or atherosclerosis). The compositions and methods may also be
XX used in assessing the histological type of neoplasm associated with
XX ovarian cancer, monitoring the progression of ovarian cancer, determining
XX whether ovarian cancer has metastasized or is likely to metastasize,
XX selecting a composition for inhibiting ovarian cancer, assessing the
XX ovarian carcinogenic potential of a compound, or inhibiting ovarian
XX cancer or at risk of developing ovarian cancer. The present amino acid
XX sequence represents one of the ovarian cancer markers described in the
XX invention
XX
SQ Sequence 243 AA;
Query Match 93.7%; Score 1225; DB 5; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.9e-115;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
QY 1 MHPOGRASPOLLLGLFLVLLLLQLSAPSSASBNPKYKOKALRQREVVLDYNGMCLQG 60
DB 1 MRPOGRASPOLRRG--LTLTLLTLLQLPAPSSASBIPKQKOKALRQREVVLDYNGMCLQG 58
QY 61 PAGVGRDSSPGANGIPGTGIPGRDGFKEGKCECLRSFESWTPNYKQCSWSSLYNGI 120
DB 59 PAEVPGRDSSPGANGIPGTGIPGRDGFKEGKCECLRSFESWTPNYKQCSWSSLYNGI 118
QY 121 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTNGAECGSPLEIAIYVL 180
DB 119 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTNGAECGSPLEIAIYVL 178
QY 181 DOGSPELNSTINIHRTSSVSGLCBEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRIT 240
DB 179 DOGSPENNSTINIHRTSSVSGLCBEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRIT 238
```

QY 241 EELPK 245  
|||  
Db 239 EELPK 243

RESULT 8  
ID AB05554  
AB05554 standard; protein; 243 AA.  
XX  
AC AB05554;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Breast cancer-associated protein 19.  
XX  
KW Breast cancer; breast cancer-associated gene sequence; drug development;  
XX pharmacogenetics; biosensor development.  
XX  
OS Unidentified.  
XX  
PN WO200259377-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 24-JAN-2002; 2002WO-US002242.  
XX  
PR 24-JAN-2001; 2001US-0263965P.  
XX 02-FEB-2001; 2001US-0265928P.  
XX 09-APR-2001; 2001US-00829472.  
XX 09-APR-2001; 2001US-0282698P.  
XX 04-MAY-2001; 2001US-0288590P.  
XX 29-MAY-2001; 2001US-0294443P.  
XX  
PA (EOSR-) EOS BIOTECHNOLOGY INC.  
XX  
PI Mack DH, Gish KC, Afar D;  
XX  
DR WPI; 2002-583738/62.  
XX N-PSDB; ABT07711.  
XX  
PT Detecting a breast cancer-associated transcript in a patient's cell,  
XX useful for diagnosing breast cancer, comprises contacting a biological  
XX sample with a polynucleotide that selectively hybridizes with breast  
XX cancer nucleic acids.  
XX  
PS Disclosure; Page 364; 414pp; English.  
XX  
XX The invention comprises a method of detecting a breast cancer-associated  
XX transcript in a cell from a patient. The method of the invention involves  
XX contacting a biological sample from the patient with a nucleotide that  
XX hybridizes to one of the 69 breast cancer-associated gene sequences shown  
XX in the specification. The method of the invention is useful in the  
XX diagnosis or prognosis of breast cancer, and for detecting genes that are  
XX up or down-regulated in breast cancer cells. Genes identified by the  
XX method of the invention can be used in diagnostic purposes and also as  
XX targets for screening for therapeutic compounds that modulate breast  
XX cancer (e.g. hormones or antibodies). Identification of genes that are  
XX over or under expressed in breast cancer can additionally provide high-  
XX resolution, high-sensitivity datasets which can be used in the areas of  
XX diagnostics, therapeutics, drug development, pharmacogenetics, protein  
XX structure and biosensor development. Amino acid sequences AB05536 -  
XX AB05604 represent the proteins encoded by the 69 breast cancer-  
XX associated genes of the invention  
XX  
SQ Sequence 243 AA.

Query Match 93.7%; Score 1225; DB 5; Length 243;  
Best Local Similarity 94.7%; Pred. No. 6.9e-115;  
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHQGRAPASPOLLLGLFVLVLLLLQLSAPSSASNPVVKAKLRQGVVDLYNMCLOG 60  
|||  
Db 1 MRQGGPAPASPOLRGG--LTLTLLLLQLPAPSSASEIRPKGKQKQLRQREVVDLYNMCLOG 58

QY 61 PAGVPGRDSSPGANGIPGRIPIGRDGFKGKGECLARESPESWTPNYKQCSWSSLNYGI 120  
|||||  
Db 59 PAGVPGRDSSPGANGIPGRIPIGRDGFKGKGECLARESPESWTPNYKQCSWSSLNYGI 118

QY 121 DLGKIAECTFTMKRNSALRVLFSGSLRLKCRNACCORWFTFNAGCSGPLPIEAIYYL 180  
|||||  
Db 119 DLGKIAECTFTMKRNSALRVLFSGSLRLKCRNACCORWFTFNAGCSGPLPIEAIYYL 178

QY 181 DQSGPELNSTINIRHTSSVEGLCEGIGAGLDVAIIVGTCSDFPYGDASTGNANSYRIRI 240  
|||||  
Db 179 DQSGPELNSTINIRHTSSVEGLCEGIGAGLDVAIIVGTCSDFPYGDASTGNANSYRIRI 238

QY 241 EELPK 245  
|||  
Db 239 EELPK 243

RESULT 9  
ID ABB80978  
ABB80978 standard; protein; 243 AA.  
XX  
AC ABB80978;  
XX  
DT 21-OCT-2002 (first entry)  
XX  
DE Human REMODELIN polypeptide.  
XX  
XX REMODELIN; vulnereary; vasotrophic; cytostatic; osteopathic; collagen;  
XX gene therapy; bone; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200242487-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 19-OCT-2001; 2001WO-US050940.  
XX  
PR 19-OCT-2000; 2000US-00692081.  
XX  
PA (MAIN-) MAINE MEDICAL CENT RES INST.  
XX  
PI Linder V, Friesel R;  
XX  
DR WPI; 2002-590472/63.  
XX N-PSDB; ABB86481.  
XX  
XX New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
XX negative remodeling, fibrosis, collagen disease, and bone disease in a  
XX mammal.  
XX  
PS Claim 5; Fig 4B; 187pp; English.  
XX  
XX The invention relates to mammalian REMODELIN polypeptides and encoding  
XX polynucleotides. REMODELIN is used to diagnose arterial restenosis,  
XX negative remodeling or fibrosis, bone disease such as osteogenesis  
XX imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis  
XX bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are  
XX used to affect cellular gene expression in a mammal, where the cellular  
XX gene is from transforming growth factor (TGF)-beta1, collagen I/II/III/IV,  
XX osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein  
XX 4. The expression of osteopontin is dependent on Cbfa1. Compositions  
XX comprising antisense REMODELIN sequences are useful for treating diseases  
XX mediated by abnormal expression of a REMODELIN molecule in a human such  
XX as impaired wound healing, fibrosis of an organ, ectopic ossification, or  
XX hypertrophic scar formation. REMODELIN is useful in the development of  
XX bone during mammalian embryogenesis, proliferation and/or migration,  
XX cellular signaling, adventitial fibrosis, negative remodeling and arterial  
XX restenosis, smooth muscle cell proliferation, and arterial remodeling and  
XX useful in gene therapy. REMODELIN is useful for elucidating the function  
XX of REMODELIN molecules in a cell, to identify a compound that affects  
XX REMODELIN expression and/or TGF-beta signaling, as a potential

CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
CC formation, to promote bone fracture healing, and to increase or decrease  
CC expression of REMODELIN in mammal. It is also useful for producing  
CC recombinant cell and transgenic non-human mammals which are useful tools  
CC for the study of REMODELIN action, for identifying novel diagnostics and  
CC therapeutics for treatment, and for elucidating the cellular roles of  
CC REMODELIN. The present sequence represents a human REMODELIN polypeptide  
XX  
SQ Sequence 243 AA;  
  
Query Match 93.7%; Score 1225; DB 5; Length 243;  
Best Local Similarity 94.7%; Pred. No. 6.9e-115;  
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;  
  
QY 1 MHPOGRAASPQQLLGLFLVLLLLQLSPSSASSENPKYKOKALIREVVDLYNGMCLQG 60  
DB 1 MRPOGPAPSPQRLRG--LTLILLLLQLPAPSSASEIPKQKQKALRQREVVDLYNGMCLQG 58  
  
QY 61 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTPNYKQCSWSLNYGI 120  
DB 59 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTPNYKQCSWSLNYGI 118  
  
QY 121 DLGKIACEFTTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIVL 180  
DB 119 DLGKIACEFTTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIVL 178  
  
QY 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYKGDASTGMNSVSRIT 240  
DB 179 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYKGDASTGMNSVSRIT 238  
  
QY 241 EELPK 245  
DB 239 EELPK 243  
  
RESULT 10  
ABR58546  
ID ABR58546 standard; protein; 243 AA.  
AC ABR58546;  
DT 09-JUL-2003 (first entry)  
XX  
DE Human cancer related protein SEQ ID NO:203.  
XX  
KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;  
XX heart disease; atherosclerosis; endometriosis.  
OS Homo sapiens.  
XX  
PN WO2003025138-A2.  
XX  
PD 27-MAR-2003.  
XX  
PF 17-SEP-2002; 2002WO-US029560.  
XX  
PR 17-SEP-2001; 2001US-0323469P.  
PR 20-SEP-2001; 2001US-0323687P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 08-FEB-2002; 2002US-0355145P.  
PR 08-FEB-2002; 2002US-0355257P.  
PR 12-APR-2002; 2002US-0372246P.  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;  
PI Zlotnick A;  
XX  
XX MPI: 2003-354600/33.  
DR N-PSDB; ACCT2666.  
XX  
PT New genes that are up-regulated or down-regulated in cancers, useful as

PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as  
PT therapeutic targets for screening drugs for treating these diseases.  
XX  
PS Claim 12; Page 736; 767pp; English.  
XX  
XX The present invention describes an isolated nucleic acid molecule, which  
CC comprises the sequence of any of the genes that are up-regulated or down-  
CC regulated in specific cancers (e.g. about 1011 genes up-regulated in  
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer  
CC related gene nucleotide sequences which encode the proteins given in  
CC ABR58521 to ABR58709. Also described: (1) determining the presence or  
CC absence of a pathological cell in a patient; (2) an expression vector  
CC comprising a nucleic acid molecule described above; (3) a host cell  
CC comprising the vector; (4) an isolated polypeptide, which is encoded by  
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide  
CC of (4); (6) specifically targeting a compound to a pathological cell in a  
CC patient by administering the nucleic acid to the patient the antibody above; and (7) a  
CC drug screening assay. The nucleic acid is useful as diagnostic markers or  
CC therapeutic targets. In particular, the nucleic acid is useful for  
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,  
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,  
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,  
CC atherosclerosis and endometriosis. The nucleic acid is also useful in  
CC drug screening, particularly for identifying agents for treating these  
CC pathologies  
XX  
SQ Sequence 243 AA;  
  
Query Match 93.7%; Score 1225; DB 6; Length 243;  
Best Local Similarity 94.7%; Pred. No. 6.9e-115;  
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;  
  
QY 1 MHPOGRAASPQQLLGLFLVLLLLQLSPSSASSENPKYKOKALIREVVDLYNGMCLQG 60  
DB 1 MRPOGPAPSPQRLRG--LTLILLLLQLPAPSSASEIPKQKQKALRQREVVDLYNGMCLQG 58  
  
QY 61 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTPNYKQCSWSLNYGI 120  
DB 59 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTPNYKQCSWSLNYGI 118  
  
QY 121 DLGKIACEFTTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIVL 180  
DB 119 DLGKIACEFTTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIVL 178  
  
QY 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYKGDASTGMNSVSRIT 240  
DB 179 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYKGDASTGMNSVSRIT 238  
  
QY 241 EELPK 245  
DB 239 EELPK 243  
  
RESULT 11  
ABR48227  
ID ABR48227 standard; protein; 243 AA.  
AC ABR48227;  
DT 12-JUN-2003 (first entry)  
XX  
DE Human bladder cancer associated protein sequence SEQ ID NO:175.  
XX  
KW Human; bladder cancer; cytostatic; gene therapy; vaccine.  
XX  
XX Homo sapiens.  
XX  
XX WO2003003906-A2.  
XX  
XX 16-JAN-2003.  
XX  
XX 03-JUL-2002; 2002WO-US021338.  
XX  
XX

PR 03-JUL-2001; 2001US-0302814P.  
PR 03-AUG-2001; 2001US-0310099P.  
PR 08-NOV-2001; 2001US-0343705P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 12-APR-2002; 2002US-0372246P.  
XX  
PA (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
PI Mack DH, Aziz N;  
XX  
DR WPI; 2003-201532/19.  
DR N-PSDB; ACC51043.  
XX  
PT Detecting a bladder cancer-associated transcript in a cell from a  
PT patient, comprises contacting a biological sample from the patient with a  
PT bladder cancer-associated polynucleotide or antibody.  
XX  
PS Claim 10; Page 289-290; 307pp; English.  
XX  
CC The present invention describes a method for detecting a bladder cancer-  
CC associated transcript in a cell from a patient. The method comprises  
CC contacting a biological sample from the patient with a polynucleotide  
CC that selectively hybridizes to a sequence that is 80 % identical to a  
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059  
CC encode the human bladder cancer-associated proteins given in ABR48146 to  
CC ABR48242). Bladder cancer-associated sequences from the present invention  
CC have cytostatic activities, and can be used in antisense gene therapy and  
CC in vaccine production. The method can be used for detecting a bladder  
CC cancer-associated transcript in a cell from a patient. The method is  
CC useful in diagnosing or treating bladder cancer and in screening for  
CC compounds that modulate bladder cancer, such as hormones or antibodies.  
CC The nucleic acid molecules from the present invention may be used in  
CC various screening and diagnostic methods, and for gene therapy, vaccine  
CC and/or antisense/inhibition applications  
XX  
SQ Sequence 243 AA;  
XX  
Query Match 93.7%; Score 1225; DB 6; Length 243;  
Best Local Similarity 94.7%; Pred. No. 6.9e-115;  
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;  
QY 1 MHPQGRASPQLLGLFLVLLLLQLSAPSSASENPKYKQKALIRQREVVDLYNMGCLQG 60  
1 MRQGPAPSPQRLRG--LTLTLTLTLQLPAPSSASEIPKQKQKQALRQREVVDLYNMGCLQG 58  
DB 119 DLGKIACCTTKRNSALRVLFSGSLRLKCRNACQRMWFTFNAGCSGPLEIAIYYL 180  
121 DLGKIACCTTKRNSALRVLFSGSLRLKCRNACQRMWFTFNAGCSGPLEIAIYYL 180  
QY 59 PAVGPRDGSPPANGIPGTPGIPGRDGFKEGKEGECLEESFEESWTPNYKQCSWSSLNNGI 118  
61 PAVGPRDGSPPANGIPGTPGIPGRDGFKEGKEGECLEESFEESWTPNYKQCSWSSLNNGI 120  
DB 119 DLGKIACCTTKRNSALRVLFSGSLRLKCRNACQRMWFTFNAGCSGPLEIAIYYL 178  
121 DLGKIACCTTKRNSALRVLFSGSLRLKCRNACQRMWFTFNAGCSGPLEIAIYYL 180  
QY 181 DQGSPELNTSTINIHRTSVGGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITII 240  
179 DQGSPELNTSTINIHRTSVGGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITII 238  
QY 241 EELPK 245  
239 EELPK 243  
DB 239 EELPK 243  
XX  
RESULT 12  
ABG75758  
ID ABG75758 standard; protein; 243 AA.  
XX  
AC ABG75758;  
XX  
DT 28-APR-2003 (first entry)  
XX  
DE Human REMODELIN protein.  
XX  
KW Human; adventitia-inducible; REMODELIN; REMODEL;

KW adventitia induced bone expressed molecule; AIBE; TGF-beta;  
KW transforming growth factor beta; adventitia; vascular remodeling;  
KW restenosis; vascular injury; antisense therapy; TGF-beta signalling;  
KW TGF-beta receptor type II; arterial remodeling; bone formation;  
KW cartilage formation; osteogenesis imperfecta; Bethlem myopathy;  
KW dystrophic epidermolysis bullosa; negative remodeling; wound healing;  
KW arterial stenosis; fibrosis; calcification; transplant;  
KW heart valve transplant; osteopathic; antiarteriosclerotic; vulnary.  
XX  
OS Homo sapiens.  
XX  
PN US2002161211-A1.  
XX  
PD 31-OCT-2002.  
XX  
PF 19-OCT-2001; 2001US-00045992.  
XX  
PR 19-OCT-2000; 2000US-00692081.  
XX  
PA (LIND/) LINDNER V.  
PA (FRIE/) FRIESEL R E.  
PI Lindner V, Friesel RE;  
XX  
DR WPI; 2003-238238/23.  
DR N-PSDB; ABX11341.  
XX  
PT New isolated REMODELIN nucleic acid and polypeptide, useful for mediating  
PT arterial remodeling, formation of bone and cartilage, and the diagnosis  
PT and treatment of disorders associated with aberrant expression of  
PT REMODELIN.  
XX  
PS Claim 3; Fig 4B; 81pp; English.  
XX  
CC The invention discloses an isolated nucleic acid encoding a mammalian  
CC adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL  
CC and/or adventitia induced bone expressed molecule, AIBE). REMODELIN  
CC expression was induced by transforming growth factor beta (TGF-beta),  
CC which is important because proliferative events occurring in the  
CC adventitia contribute to vascular remodeling and restenosis in response  
CC to vascular injury and TGF-beta has been shown to be a factor involved in  
CC this. Also disclosed is an antibody raised against REMODELIN and methods  
CC for treating a disease mediated by abnormal expression of a REMODELIN in  
CC a human, for identifying a compound that affects or reduces expression of  
CC REMODELIN in a cell (e.g. antisense therapy), for identifying a compound  
CC that affects TGF-beta signalling and for increasing or reducing REMODELIN  
CC expression in a mammal, comprising administering a REMODELIN expression  
CC increasing or reducing TGF-beta to the mammal, thereby increasing  
CC REMODELIN expression or inhibiting signalling via the TGF-beta receptor  
CC type II and reducing expression of REMODELIN in the mammal. The methods  
CC and compositions of the present invention are useful for mediating  
CC arterial remodeling, formation of bone and cartilage and the diagnosis  
CC and treatment of disorders associated with aberrant expression of  
CC REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis  
CC bullosa, Bethlem myopathy, negative remodeling, wound healing, arterial  
CC stenosis, vessel injury, fibrosis and calcification of a transplant,  
CC preferably a heart valve transplant. The sequence presented is the human  
CC REMODELIN protein  
XX  
SQ Sequence 243 AA;  
XX  
Query Match 93.7%; Score 1225; DB 6; Length 243;  
Best Local Similarity 94.7%; Pred. No. 6.9e-115;  
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;  
QY 1 MHPQGRASPQLLGLFLVLLLLQLSAPSSASENPKYKQKALIRQREVVDLYNMGCLQG 60  
1 MRQGPAPSPQRLRG--LTLTLTLTLQLPAPSSASEIPKQKQKQALRQREVVDLYNMGCLQG 58  
DB 119 DLGKIACCTTKRNSALRVLFSGSLRLKCRNACQRMWFTFNAGCSGPLEIAIYYL 180  
121 DLGKIACCTTKRNSALRVLFSGSLRLKCRNACQRMWFTFNAGCSGPLEIAIYYL 180  
QY 59 PAVGPRDGSPPANGIPGTPGIPGRDGFKEGKEGECLEESFEESWTPNYKQCSWSSLNNGI 118  
61 PAVGPRDGSPPANGIPGTPGIPGRDGFKEGKEGECLEESFEESWTPNYKQCSWSSLNNGI 120  
DB 119 DLGKIACCTTKRNSALRVLFSGSLRLKCRNACQRMWFTFNAGCSGPLEIAIYYL 178  
121 DLGKIACCTTKRNSALRVLFSGSLRLKCRNACQRMWFTFNAGCSGPLEIAIYYL 180

QY 121 DLGKIAECTFTKRSNSALRVLPFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIITL 180  
 DB 119 DLGKIAECTFTKRSNSALRVLPFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIITL 178  
 QY 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYPKGDASTGMNSVSRITL 240  
 DB 179 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYPKGDASTGMNSVSRITL 238  
 QY 241 EELPK 245  
 DB 239 EELPK 243  
 RESULT 13  
 ID ABUS6607 standard; protein; 243 AA.  
 AC ABUS6607;  
 XX  
 DT 02-APR-2003 (first entry)  
 XX  
 DE Lung cancer-associated polypeptide #200.  
 XX  
 KM Lung cancer-associated polypeptide; cytosstatic; emphysema;  
 KM antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
 KM small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KM chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KM interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
 XX  
 OS Unidentified.  
 OS  
 PN WO200286443-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 18-APR-2002; 2002WO-US012476.  
 XX  
 PR 18-APR-2001; 2001US-0284770P.  
 PR 10-MAY-2001; 2001US-0290492P.  
 PR 09-NOV-2001; 2001US-0339245P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 29-NOV-2001; 2001US-0334370P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Aziz N, Murray R;  
 XX  
 DR WPI: 2003-093161/08.  
 DR N-PSDB; ABR76336.  
 XX  
 PT Detecting a lung cancer-associated transcript in a cell from a patient  
 PT for treating lung cancer, by contacting a biological sample from the  
 PT patient with a polynucleotide that exhibits increased or decreased  
 PT expression in lung cancer.  
 XX  
 PS Claim 27; Page 337; 453pp; English.  
 XX  
 CC The invention relates to a method for detecting a lung cancer-associated  
 CC transcript in a cell from a patient, comprising contacting a biological  
 CC sample from the patient with a polynucleotide that selectively hybridizes  
 CC to a sequence that is at least 80 % identical to a gene that exhibits  
 CC increased or decreased expression in lung cancer samples. Lung cancer-  
 CC associated polynucleotides and polypeptides are used for identifying a  
 CC compound that modulates a lung cancer-associated polypeptide, for  
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
 CC cancer in a patient and for treating a mammal having lung cancer by  
 CC administering a modulatory compound identified. The methods are useful  
 CC for treating lung cancer, such as small cell lung cancer, non-small cell  
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful

CC for diagnostic purposes and as targets for screening for therapeutic  
 CC compounds that modulate lung cancer, such as antibodies. Sequences  
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the  
 CC invention  
 XX  
 SO Sequence 243 AA;  
 Query Match 93.7%; Score 1225; DB 6; Length 243;  
 Best Local Similarity 94.7%; Pred. No. 6; 9e-115;  
 Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;  
 QY 1 MHPQGRASPOLLILGLVLLILOLSAPSSASENPKVOKALIROREVVLDYNGWCLQG 60  
 DB 1 MHPQGRASPORLRC--LILLILLQLPAPSSASEIPKQKOKLROREVVLDYNGWCLQG 58  
 QY 61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGCECLRESFEBSWTPNYKQCSWSLNYGI 120  
 DB 59 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGCECLRESFEBSWTPNYKQCSWSLNYGI 118  
 QY 121 DLGKIAECTFTKRSNSALRVLPFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIITL 180  
 DB 119 DLGKIAECTFTKRSNSALRVLPFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIITL 178  
 QY 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYPKGDASTGMNSVSRITL 240  
 DB 179 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYPKGDASTGMNSVSRITL 238  
 QY 241 EELPK 245  
 DB 239 EELPK 243  
 RESULT 14  
 ID ADB80510  
 AC ADB80510 standard; protein; 243 AA.  
 XX  
 AC ADB80510;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Ovarian cancer-associated protein #37.  
 XX  
 KM cytosstatic; gene therapy; vaccine; ovarian cancer; diagnosis;  
 KM post-operative chemotherapy; radiation therapy; tumour prognosis;  
 KM pre-cancerous lesion detection.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO2002102235-A2.  
 XX  
 PD 27-DEC-2002.  
 XX  
 PF 18-JUN-2002; 2002WO-US019297.  
 XX  
 PR 18-JUN-2001; 2001US-0299234P.  
 PR 27-AUG-2001; 2001US-0315287P.  
 PR 05-SEP-2001; 2001US-0317544P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Mack DH, Gish KC;  
 XX  
 DR WPI: 2003-167431/16.  
 DR N-PSDB; ADB80509.  
 XX  
 PT Detecting an ovarian cancer-associated transcript in a cell from a  
 PT patient, comprising contacting a biological sample from the patient with a  
 PT polynucleotide that hybridizes to an ovarian cancer gene.  
 XX  
 PS Claim 13; Page 299; 332pp; English.

CC The invention relates to a method of detecting an ovarian cancer-associated transcript in a cell from a patient, by contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to any of one of 80 nucleic acid sequences given in the specification. The method is useful in diagnosing ovarian cancer and in identifying and using agents and/or targets that inhibit ovarian cancer. The nucleic acid molecule, CC polypeptide and the antibody may also be used in detecting ovarian cancers, monitoring and early detection of relapse following treatment, CC monitoring response to therapy, selecting patients for post-operative chemotherapy or radiation therapy, in selecting mode of therapy, CC determining tumour prognosis, early detection of pre-cancerous lesions, CC and as vaccines. This sequence corresponds to one of the proteins used for the detection method of the invention.

XX Sequence 243 AA;

Query Match 93.7%; Score 1225; DB 7; Length 243;

Best Local Similarity 94.7%; Pred. No. 6.9e-115;

Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPQGRASPOLLLGLFLVLLLLLSAPSSASENPKYKOKALIRQREVVDLYNMCLOG 60  
 DB 1 MHPQGRASPOLRLRG--LILLILLQLPAPSSASEIPKQKQALRQREVVDLYNMCLOG 58  
 QY 61 PAVGPRDSSPGANGIPGTGIPGRDGFKEGECCLRESFEESMTPNYKQCSWSLNYGI 120  
 DB 59 PAVGPRDSSPGANGIPGTGIPGRDGFKEGECCLRESFEESMTPNYKQCSWSLNYGI 118  
 QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKRNACCORWYFTFNAGCSGPLPIEAI1YL 180  
 DB 119 DLGKIACFTTKRNSALRVLFSGSLRLKRNACCORWYFTFNAGCSGPLPIEAI1YL 178  
 QY 181 DOGSEPLNSTINIHRSSVVGCLCEGIGAGLVDAIIVGTCSYPRKGDASTGMNSVSR11I 240  
 DB 179 DOGSEPLNSTINIHRSSVVGCLCEGIGAGLVDAIIVGTCSYPRKGDASTGMNSVSR11I 238  
 QY 241 EELPK 245  
 DB 239 EELPK 243

RESULT 15

ADN39855 ID ADN39855 standard; protein; 243 AA.

AC ADN39855;

DT 17-JUN-2004 (first entry)

DE Cancer/angiogenesis/fibrois-related polypeptide, SEQ ID NO:C225.

XX Human; differential expression; cancer; angiogenic disorder;  
 KM fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
 KM inflammatory disease; autoimmune disease;  
 KM retinal neovascularisation syndrome; scarring; uterine fibroid;  
 KM detection; diagnosis; prognosis; drug screening; drug targeting;  
 KM wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
 KM vulneraty; gene therapy; vaccine.

XX Homo sapiens.

PN WO2003042661-A2.

PD 22-MAY-2003.

PF 13-NOV-2002; 2002WO-US036810.

PR 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.  
 PR 10-JAN-2002; 2002US-0347349P.  
 PR 08-FEB-2002; 2002US-0355250P.  
 PR 13-FEB-2002; 2002US-0356714P.  
 PR 20-FEB-2002; 2002US-0359077P.  
 PR 29-MAR-2002; 2002US-0368809P.  
 PR 04-APR-2002; 2002US-0370110P.  
 PR 12-APR-2002; 2002US-0372246P.  
 PR 05-JUN-2002; 2002US-0386614P.  
 PR 16-JUL-2002; 2002US-0396839P.  
 PR 22-JUL-2002; 2002US-0397775P.  
 PR 22-JUL-2002; 2002US-0397845P.  
 PR 09-SEP-2002; 2002US-0409450P.

XX (EOS-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glyme R, Hevazi PA;

PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX WPI: 2003-468649/44.

DR N-PSDB; ADN39637.

XX Determining the presence or absence of a pathological cell in a patient, useful for diagnosing, prognosing or treating cancer, comprises detecting a nucleic acid in a biological sample.

PS Claim 12; SEQ ID NO C225; 1385pp; English.

CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; CC and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, CC antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal CC neovascularisation syndromes, scarring and uterine fibroids. They may CC also be useful in wound healing and in contraception. The present CC sequence represents a polypeptide of the invention.

XX Sequence 243 AA;

Query Match 93.7%; Score 1225; DB 7; Length 243;

Best Local Similarity 94.7%; Pred. No. 6.9e-115;

Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPQGRASPOLLLGLFLVLLLLLSAPSSASENPKYKOKALIRQREVVDLYNMCLOG 60  
 DB 1 MHPQGRASPOLRLRG--LILLILLQLPAPSSASEIPKQKQALRQREVVDLYNMCLOG 58  
 QY 61 PAVGPRDSSPGANGIPGTGIPGRDGFKEGECCLRESFEESMTPNYKQCSWSLNYGI 120  
 DB 59 PAVGPRDSSPGANGIPGTGIPGRDGFKEGECCLRESFEESMTPNYKQCSWSLNYGI 118  
 QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKRNACCORWYFTFNAGCSGPLPIEAI1YL 180  
 DB 119 DLGKIACFTTKRNSALRVLFSGSLRLKRNACCORWYFTFNAGCSGPLPIEAI1YL 178  
 QY 181 DOGSEPLNSTINIHRSSVVGCLCEGIGAGLVDAIIVGTCSYPRKGDASTGMNSVSR11I 240  
 DB 179 DOGSEPLNSTINIHRSSVVGCLCEGIGAGLVDAIIVGTCSYPRKGDASTGMNSVSR11I 238  
 QY 241 EELPK 245  
 DB 239 EELPK 243

Thu Aug 4 05:41:41 2005

us-10-634-108-2.rag

Page 11

Search completed: August 4, 2005, 04:16:36  
Job time : 98.4771 secs

---

**This Page Blank (uspis)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 4, 2005, 04:21:44 ; Search time 86.7909 Seconds  
(without alignments)  
1099.816 Million cell updates/sec

Title: US-10-634-108-2

Perfect score: 1307  
Sequence: 1 MHPOGRRAAPOLLGLGLFLVL.....GDASTGMSVSRITIELEK 245

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1745140 seqs, 389608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Published Applications AA:  
1: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/2/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/2/pubppa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
18: /cgn2\_6/prodata/2/pubppa/US10\_NEW\_PUB.pep:\*  
19: /cgn2\_6/prodata/2/pubppa/US11\_PUBCOMB.pep:\*  
20: /cgn2\_6/prodata/2/pubppa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/prodata/2/pubppa/US60\_NEW\_PUB.pep:\*  
22: /cgn2\_6/prodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1307	100.0	245	US-10-045-992-2	Sequence 2, Appl1
2	1307	100.0	245	US-10-634-108-2	Sequence 2, Appl1
3	1307	100.0	245	US-10-939-233-2	Sequence 2, Appl1
4	1307	100.0	277	US-10-045-992-5	Sequence 5, Appl1
5	1307	100.0	277	US-10-634-108-5	Sequence 5, Appl1
6	1307	100.0	277	US-10-939-233-5	Sequence 5, Appl1
7	1225	93.7	243	US-09-829-472A-19	Sequence 19, Appl1
8	1225	93.7	243	US-10-045-992-4	Sequence 4, Appl1
9	1225	93.7	243	US-10-097-340-135	Sequence 135, App
10	1225	93.7	243	US-10-295-027-50	Sequence 50, Appl
11	1225	93.7	243	US-10-295-027-1173	Sequence 1173, Ap

12	1225	93.7	243	US-10-173-999-74	Sequence 74, Appl
13	1225	93.7	243	US-10-058-270A-38	Sequence 38, Appl
14	1225	93.7	243	US-10-634-108-4	Sequence 4, Appl1
15	1225	93.7	243	US-10-188-832-175	Sequence 175, App
16	1225	93.7	243	US-10-939-233-4	Sequence 4, Appl1
17	1225	93.7	278	US-10-097-340-131	Sequence 131, Appl
18	1225	93.7	278	US-10-177-293-496	Sequence 496, App
19	1225	93.7	278	US-10-301-822-77	Sequence 77, Appl
20	1225	93.7	278	US-10-296-115-1261	Sequence 1261, Ap
21	1225	93.7	278	US-10-961-139-2	Sequence 2, Appl1
22	1222	93.5	243	US-10-351-334-205	Sequence 205, App
23	1222	93.5	278	US-10-060-036-4551	Sequence 4551, App
24	1220	93.3	243	US-10-176-847-14	Sequence 14, Appl
25	1216	93.0	243	US-09-834-759-514	Sequence 514, App
26	1216	93.0	243	US-09-938-418-7	Sequence 7, Appl1
27	1216	93.0	243	US-09-946-374-352	Sequence 352, App
28	1216	93.0	243	US-09-946-374-431	Sequence 431, App
29	1216	93.0	243	US-10-006-867-122	Sequence 122, App
30	1216	93.0	243	US-10-052-566-366	Sequence 366, App
31	1216	93.0	243	US-10-007-805-514	Sequence 514, App
32	1216	93.0	243	US-10-063-547-122	Sequence 122, App
33	1216	93.0	243	US-10-063-551-122	Sequence 122, App
34	1216	93.0	243	US-10-174-590-366	Sequence 366, App
35	1216	93.0	243	US-10-176-758-366	Sequence 366, App
36	1216	93.0	243	US-10-175-737-366	Sequence 366, App
37	1216	93.0	243	US-10-063-616-122	Sequence 122, App
38	1216	93.0	243	US-10-174-581-366	Sequence 366, App
39	1216	93.0	243	US-10-176-483-366	Sequence 366, App
40	1216	93.0	243	US-10-176-749-366	Sequence 366, App
41	1216	93.0	243	US-10-176-914-366	Sequence 366, App
42	1216	93.0	243	US-10-176-915-366	Sequence 366, App
43	1216	93.0	243	US-10-063-553-122	Sequence 122, App
44	1216	93.0	243	US-10-063-513-122	Sequence 122, App
45	1216	93.0	243	US-10-063-515-122	Sequence 122, App

## ALIGNMENTS

RESULT 1  
US-10-045-992-2 ; Application US/10045992  
; Publication No. US20020161211A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volhard  
; APPLICANT: FRIEDEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN  
; FILE REFERENCE: 053689-5006-01  
; CURRENT APPLICATION NUMBER: US/10/045,992  
; CURRENT FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-045-992-2

Query Match 100.0%; Score 1307; DB 13; Length 245;  
Best Local Similarity 100.0%; Pred. No. 4.7e-121;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MHPOGRRAAPOLLGLGLFLVLILLOLSASSASSENPKVKOKALIRQREVVDLYNGKCLQG 60  
DB 1 MHPOGRRAAPOLLGLGLFLVLILLOLSASSASSENPKVKOKALIRQREVVDLYNGKCLQG 60  
QY 61 PAVGVRGDDSPRANGIRPTGIRPGRDGFKEGEGECIRBSFESWTPNYQCWSLNYGI 120  
DB 61 PAVGVRGDDSPRANGIRPTGIRPGRDGFKEGEGECIRBSFESWTPNYQCWSLNYGI 120  
QY 121 DVGKIAECTFTMRNSALRVLFSGSLRLKCNACCQQRVYFFPFGNAGCGPIFEATIVL 180

```
Db      121 |||||
        DGGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTNGAECGSLPIEAIITL 180
Qy      181 DQSSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIIT 240
        |||||
        DQSSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIIT 240
Db      241 EELPK 245
        |||||
        241 EELPK 245

RESULT 2
US-10-634-108-2
; Sequence 2, Application US/10634108
; Publication No. US2004063176A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/10/634,108
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 245
; ORGANISM: Rattus sp.
US-10-634-108-2

Query Match      100.0%; Score 1307; DB 15; Length 245;
Best Local Similarity 100.0%; Pred. No. 4,7e-121;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MHPOGRAASPOLLGFLVTLILLLOLSAPSSASENPKVKOKALIROREVVDLYNGMCLQG 60
        1 MHPOGRAASPOLLGFLVTLILLLOLSAPSSASENPKVKOKALIROREVVDLYNGMCLQG 60
Db      61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSINTYGI 120
        61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSINTYGI 120
Db      121 DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTNGAECGSLPIEAIITL 180
        |||||
        DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTNGAECGSLPIEAIITL 180
Qy      181 DQSSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIIT 240
        |||||
        DQSSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIIT 240
Db      241 EELPK 245
        |||||
        241 EELPK 245

RESULT 3
US-10-939-233-2
; Sequence 2, Application US/10939233
; Publication No. US20050147602A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert E.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTRHCL, A NOVEL
; FILE REFERENCE: 053689-5006-03
; CURRENT APPLICATION NUMBER: US/10/939,233
; PRIOR FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 09/692,081
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US 10/045,992
```

```
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-939-233-2
```

```
Query Match      100.0%; Score 1307; DB 18; Length 245;
Best Local Similarity 100.0%; Pred. No. 4,7e-121;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MHPOGRAASPOLLGFLVTLILLLOLSAPSSASENPKVKOKALIROREVVDLYNGMCLQG 60
        1 MHPOGRAASPOLLGFLVTLILLLOLSAPSSASENPKVKOKALIROREVVDLYNGMCLQG 60
Db      61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSINTYGI 120
        61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSINTYGI 120
Db      121 DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTNGAECGSLPIEAIITL 180
        |||||
        DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTNGAECGSLPIEAIITL 180
Qy      181 DQSSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIIT 240
        |||||
        DQSSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIIT 240
Db      241 EELPK 245
        |||||
        241 EELPK 245
```

```
RESULT 4
US-10-045-992-5
; Sequence 5, Application US/10045992
; Publication No. US20020161211A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert E.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN
; FILE REFERENCE: 053689-5006-01
; CURRENT APPLICATION NUMBER: US/10/045,992
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-045-992-5

Query Match      100.0%; Score 1307; DB 13; Length 277;
Best Local Similarity 100.0%; Pred. No. 5,5e-121;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MHPOGRAASPOLLGFLVTLILLLOLSAPSSASENPKVKOKALIROREVVDLYNGMCLQG 60
        33 MHPOGRAASPOLLGFLVTLILLLOLSAPSSASENPKVKOKALIROREVVDLYNGMCLQG 92
Db      61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSINTYGI 120
        93 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSINTYGI 152
Db      121 DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTNGAECGSLPIEAIITL 180
        153 DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWYFTNGAECGSLPIEAIITL 212
Qy      181 DQSSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIIT 240
        |||||
```

Db 213 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIII 272  
QY 241 EELPK 245  
273 EELPK 277  
RESULT 5  
US-10-634-108-5  
; Sequence 5, Application US/10634108  
; Publication No. US20040063176A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volhard  
; APPLICANT: FRIESEL, Robert F.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL  
; FILE REFERENCE: 36-13 Lindner et al. (20036.0013)  
; CURRENT APPLICATION NUMBER: US/10/634,108  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-634-108-5

Query Match 100.0%; Score 1307; DB 15; Length 277;  
Best Local Similarity 100.0%; Pred. No. 5, 5e-121;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPOGRAASPOLLIGFLVLLILLQISAPSSASENPKVKOKALIRORREVVLDYNGMCLOG 60  
Db 33 MHPOGRAASPOLLIGFLVLLILLQISAPSSASENPKVKOKALIRORREVVLDYNGMCLOG 92  
QY 61 PAVPGRDSSPGANGIPGTGIPGRDGFGEKGECLRESFEESWTPNYQCWSLSIYGI 120  
Db 93 PAVPGRDSSPGANGIPGTGIPGRDGFGEKGECLRESFEESWTPNYQCWSLSIYGI 152  
QY 121 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIIVL 180  
Db 153 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIIVL 212  
QY 181 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIII 240  
Db 213 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIII 272  
QY 241 EELPK 245  
Db 273 EELPK 277  
RESULT 6  
US-10-939-233-5  
; Sequence 5, Application US/10939233  
; Publication No. US20050147602A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volhard  
; APPLICANT: FRIESEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTRHC1, A NOVEL  
; TITLE OF INVENTION: MODULATOR OF COLLAGEN MATRIX  
; FILE REFERENCE: 053689-5006-03  
; CURRENT APPLICATION NUMBER: US/10/939,233  
; CURRENT FILING DATE: 2004-09-10  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US 10/045,992  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 277

; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-939-233-5

Query Match 100.0%; Score 1307; DB 18; Length 277;  
Best Local Similarity 100.0%; Pred. No. 5, 5e-121;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPOGRAASPOLLIGFLVLLILLQISAPSSASENPKVKOKALIRORREVVLDYNGMCLOG 60  
Db 33 MHPOGRAASPOLLIGFLVLLILLQISAPSSASENPKVKOKALIRORREVVLDYNGMCLOG 92  
QY 61 PAVPGRDSSPGANGIPGTGIPGRDGFGEKGECLRESFEESWTPNYQCWSLSIYGI 120  
Db 93 PAVPGRDSSPGANGIPGTGIPGRDGFGEKGECLRESFEESWTPNYQCWSLSIYGI 152  
QY 121 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIIVL 180  
Db 153 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIIVL 212  
QY 181 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIII 240  
Db 213 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIII 272  
QY 241 EELPK 245  
Db 273 EELPK 277

RESULT 7  
US-09-829-472A-19  
; Sequence 19, Application US/09829472A  
; Publication No. US20040146862A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David  
; APPLICANT: Gish, Kurt  
; APPLICANT: EOS Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and  
; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer  
; FILE REFERENCE: 018501-001200US  
; CURRENT APPLICATION NUMBER: US/09/829,472A  
; CURRENT FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 09/525,361  
; PRIOR FILING DATE: 2000-03-15  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: BCN4, ESTs, secreted protein  
US-09-829-472A-19

Query Match 93.7%; Score 1225; DB 12; Length 243;  
Best Local Similarity 94.7%; Pred. No. 6, 3e-113;  
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPOGRAASPOLLIGFLVLLILLQISAPSSASENPKVKOKALIRORREVVLDYNGMCLOG 60  
Db 1 MHPOGRAASPOLLIGFLVLLILLQISAPSSASENPKVKOKALIRORREVVLDYNGMCLOG 58  
QY 61 PAVPGRDSSPGANGIPGTGIPGRDGFGEKGECLRESFEESWTPNYQCWSLSIYGI 120  
Db 59 PAVPGRDSSPGANGIPGTGIPGRDGFGEKGECLRESFEESWTPNYQCWSLSIYGI 118  
QY 121 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIIVL 180  
Db 119 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIIVL 178  
QY 181 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIII 240  
Db 179 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIII 238



```
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-50
```

```
Query Match 93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 222; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```
QY 1 MHQGRASAPQLLGLFVLLLLQLSAPSSASENPKYKQKALIRQREVVDLYNMCLOG 60
DB 1 MRQGPASAPQRLRG--LTLTLLLLQLPAPSSASBEPKQKQKQLRQREVVDLYNMCLOG 58
QY 61 PAVPGRDSSPGANGIPGTGPIGRDGFGEKGECECLRESPEESWTPNYKQCSMSLNYGI 120
DB 59 PAVPGRDSSPGANGIPGTGPIGRDGFGEKGECECLRESPEESWTPNYKQCSMSLNYGI 118
QY 121 DGLKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNGAECGSPPIEAIIVL 180
DB 119 DGLKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNGAECGSPPIEAIIVL 178
QY 181 DQSPPELNSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNVSRIII 240
DB 179 DQSPPELNSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNVSRIII 238
QY 241 EELPK 245
DB 239 EELPK 243
```

## RESULT 11

```
US-10-295-027-1173
; Sequence 1173, Application US/10295027
; Publication No. US2003022350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natascha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyme, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
```

```
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1173
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1173
```

```
Query Match 93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 222; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```
QY 1 MHQGRASAPQLLGLFVLLLLQLSAPSSASENPKYKQKALIRQREVVDLYNMCLOG 60
DB 1 MRQGPASAPQRLRG--LTLTLLLLQLPAPSSASBEPKQKQKQLRQREVVDLYNMCLOG 58
QY 61 PAVPGRDSSPGANGIPGTGPIGRDGFGEKGECECLRESPEESWTPNYKQCSMSLNYGI 120
DB 59 PAVPGRDSSPGANGIPGTGPIGRDGFGEKGECECLRESPEESWTPNYKQCSMSLNYGI 118
QY 121 DGLKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNGAECGSPPIEAIIVL 180
DB 119 DGLKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNGAECGSPPIEAIIVL 178
QY 181 DQSPPELNSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNVSRIII 240
DB 179 DQSPPELNSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNVSRIII 238
QY 241 EELPK 245
DB 239 EELPK 243
```

## RESULT 12

```
US-10-173-999-74
; Sequence 74, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
```

```
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-74

Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPOGRAASPQULLGLFLVLLLLQLSAPSSASENPKYKOKALIROREVVDLYNMCLOG 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRPOGPAAAPQRLRG--LTLTLLTLLQLPAPSSASEIPKQKQKQLRQREVVDLYNMCLOG 58
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PAVPGRDGPANGIPGTGIPGRDGFKEGEGECLRESFESWTPTYKQCSWSSLNTGI 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 59 PAVPGRDGPANGIPGTGIPGRDGFKEGEGECLRESFESWTPTYKQCSWSSLNTGI 118
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 DLGKIACFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPLPIEAIITL 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 119 DLGKIACFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPLPIEAIITL 178
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 DOGSEPLNSTINIHTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRIT 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 179 DOGSEPLNSTINIHTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRIT 238
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 EELPK 245
    |||||
Db 239 EELPK 243
    |||||

RESULT 13
US-10-058-270A-38
; Sequence 38, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Afar, Daniel
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer Compositions and
; FILE REFERENCE: 018501-005210US
; CURRENT APPLICATION NUMBER: US/10/058,270A
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 60/263,965
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 60/265,928
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 09/829,472
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/282,698
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/288,590
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,443
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-270A-38

Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```
QY 1 MHPOGRAASPQULLGLFLVLLLLQLSAPSSASENPKYKOKALIROREVVDLYNMCLOG 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRPOGPAAAPQRLRG--LTLTLLTLLQLPAPSSASEIPKQKQKQLRQREVVDLYNMCLOG 58
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PAVPGRDGPANGIPGTGIPGRDGFKEGEGECLRESFESWTPTYKQCSWSSLNTGI 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 59 PAVPGRDGPANGIPGTGIPGRDGFKEGEGECLRESFESWTPTYKQCSWSSLNTGI 118
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 DLGKIACFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPLPIEAIITL 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 119 DLGKIACFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPLPIEAIITL 178
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 DOGSEPLNSTINIHTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRIT 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 179 DOGSEPLNSTINIHTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRIT 238
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 EELPK 245
    |||||
Db 239 EELPK 243
    |||||

RESULT 14
US-10-634-108-4
; Sequence 4, Application US/10634108
; Publication No. US20040063176A1
; GENERAL INFORMATION:
; APPLICANT: LINDER, Volhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Linder et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/10/634,108
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-108-4

Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPOGRAASPQULLGLFLVLLLLQLSAPSSASENPKYKOKALIROREVVDLYNMCLOG 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRPOGPAAAPQRLRG--LTLTLLTLLQLPAPSSASEIPKQKQKQLRQREVVDLYNMCLOG 58
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PAVPGRDGPANGIPGTGIPGRDGFKEGEGECLRESFESWTPTYKQCSWSSLNTGI 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 59 PAVPGRDGPANGIPGTGIPGRDGFKEGEGECLRESFESWTPTYKQCSWSSLNTGI 118
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 DLGKIACFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPLPIEAIITL 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 119 DLGKIACFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPLPIEAIITL 178
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 DOGSEPLNSTINIHTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRIT 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 179 DOGSEPLNSTINIHTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRIT 238
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 EELPK 245
    |||||
Db 239 EELPK 243
    |||||

RESULT 15
US-10-188-832-175
; Sequence 175, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natesha
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 175
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-175
```

```

Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```

QY      1 MHPQGAASPOLLLGLFVLLILLQLSAPSSASENPKVKALIRQREVVDLYNGMCLQG 60
      1 MRPQGPASPQRLRG--LTLILLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQG 58
QY      61 PAGVPERDGSFGANGIPGTFPGIPGRDGFKEGKEGCELRSPFESMTPNYKQCSWSLNYGI 120
      59 PAGVPERDGSFGANGIPGTFPGIPGRDGFKEGKEGCELRSPFESMTPNYKQCSWSLNYGI 118
Db
QY      121 DLGKIAECTFTTKRNSNSALRVLFSSSLRLKCRNACCQRMWTFPNGAECGGLPIEAIIYL 180
      119 DLGKIAECTFTTKRNSNSALRVLFSSSLRLKCRNACCQRMWTFPNGAECGGLPIEAIIYL 178
Db
QY      181 DQSSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGWNVSRIIIT 240
      179 DQSSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGWNVSRIIIT 238
Db
QY      241 EELPK 245
      239 EELPK 243
Db
```

```

Search completed: August 4, 2005, 04:45:26
Job time : 89.7909 secs
```

This Page Blank (uspio)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:56 ; Search time 27.881 Seconds

(Without alignments)  
741.643 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476

Sequence: 1 MRPAELGQTLSPRAGLCRPL.....GDASTGMNSVSRITIEELPK 277

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/aa/5A.COMB.pep.\*  
2: /cgn2\_6/prodata/1/aa/5B.COMB.pep.\*  
3: /cgn2\_6/prodata/1/aa/6A.COMB.pep.\*  
4: /cgn2\_6/prodata/1/aa/6B.COMB.pep.\*  
5: /cgn2\_6/prodata/1/aa/PCUS.COMB.pep.\*  
6: /cgn2\_6/prodata/1/aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1476	100.0	277	4	US-09-692-081-5 Sequence 5, Appl1
2	1307	88.6	245	4	US-09-692-081-2 Sequence 2, Appl1
3	1234.5	83.6	278	4	US-09-834-759-515 Sequence 515, App
4	1225	83.0	243	4	US-09-692-081-4 Sequence 4, Appl1
5	1222	82.8	243	4	US-09-489-847-205 Sequence 205, App
6	1216	82.4	243	4	US-09-834-759-514 Sequence 514, App
7	1103.5	74.8	232	4	US-09-205-258-958 Sequence 517, App
8	986.5	66.8	232	4	US-09-834-759-517 Sequence 516, App
9	968	65.6	197	4	US-09-834-759-516 Sequence 962, App
10	353	23.9	66	4	US-09-205-258-962 Sequence 961, App
11	296	20.1	52	4	US-09-205-258-961 Sequence 963, App
12	270	18.3	51	4	US-09-205-258-963 Sequence 960, App
13	248	16.8	46	4	US-09-834-759-518 Sequence 519, App
14	215	14.6	26	4	US-09-205-258-960 Sequence 519, App
15	141	9.6	52	4	US-09-834-759-519 Sequence 240, App
16	130	8.8	93	4	US-09-489-847-240 Sequence 362, App
17	130	8.8	93	4	US-09-489-847-362 Sequence 362, App
18	125	8.5	623	3	US-09-029-348-3 Sequence 3, Appl1
19	125	8.5	626	3	US-09-029-348-2 Sequence 2, Appl1
20	124.5	8.4	1516	4	US-09-949-016-8209 Sequence 8209, Ap
21	121.5	8.2	128	3	US-09-227-357-190 Sequence 190, Appl
22	121	8.2	755	4	US-09-919-497-57 Sequence 57, Appl1
23	119.5	8.1	246	2	US-08-463-911-4 Sequence 4, Appl1
24	119.5	8.1	313	4	US-09-949-016-9265 Sequence 9265, Ap
25	119.5	8.1	1670	4	US-09-949-016-5883 Sequence 5883, Ap
26	119	8.1	245	4	US-09-552-225A-4 Sequence 4, Appl1
27	119	8.1	245	4	US-09-552-204A-4 Sequence 4, Appl1

28	117	7.9	96	4	US-09-513-999C-4206 Sequence 4206, Ap
29	117	7.9	96	4	US-09-471-276-823 Sequence 823, App
30	117	7.9	208	4	US-09-800-729-151 Sequence 151, App
31	117	7.9	1057	3	US-08-931-820-4 Sequence 4, Appl1
32	117	7.9	1078	3	US-08-963-825-21 Sequence 21, Appl1
33	117	7.9	1078	3	US-09-500-811-21 Sequence 21, Appl1
34	117	7.9	1078	3	US-09-570-573-21 Sequence 21, Appl1
35	117	7.9	1078	3	US-09-548-608-21 Sequence 21, Appl1
36	116	7.9	258	4	US-09-976-594-815 Sequence 815, App
37	116	7.9	744	4	US-09-949-016-9607 Sequence 9607, App
38	115.5	7.8	349	6	US-10-466-2 Patent No. 5510466
39	115.5	7.8	349	6	US-10-466-2 Patent No. 5510466
40	115.5	7.8	453	6	US-10-466-4 Patent No. 5510466
41	115.5	7.8	453	6	US-10-466-4 Patent No. 5510466
42	115.5	7.8	684	1	US-08-555-669-12 Sequence 12, Appl1
43	115.5	7.8	684	1	US-09-073-663-12 Sequence 12, Appl1
44	115	7.8	1078	4	US-09-949-016-1185 Sequence 1185, A
45	115	7.8	1143	4	US-09-949-016-6137 Sequence 6137, Ap

## ALIGNMENTS

```
RESULT 1
US-09-692-081-5
; Sequence 5, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al (2000036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-692-081-5

Query Match          100.0%; Score 1476; DB 4; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.3e-150;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPAELGQTLSPRAGLCRPLCLLCSQULPHTMHPGGRASPQLLGLFLVLLLLQLSA 60
      |||
Db       1 MRPAELGQTLSPRAGLCRPLCLLCSQULPHTMHPGGRASPQLLGLFLVLLLLQLSA 60
QY      61 PSSASNPVKQKALIRORREYVDLYNGMCLQGPAGVPRDGSFGANGIRGTPTGIPRDRG 120
      |||
Db       61 PSSASNPVKQKALIRORREYVDLYNGMCLQGPAGVPRDGSFGANGIRGTPTGIPRDRG 120
QY      121 KGEKGECLRESFEESFTPNYKOCMSGLNYGIDLGIACTPTPKMSNSALRVLPFGSLR 180
      |||
Db       121 KGEKGECLRESFEESFTPNYKOCMSGLNYGIDLGIACTPTPKMSNSALRVLPFGSLR 180
QY      181 LKCRNACCQRMWFTFNGAECGSLPIEATIIYLDQSPPELINSTINIHRTSVVEGLCEGIGA 240
      |||
Db       181 LKCRNACCQRMWFTFNGAECGSLPIEATIIYLDQSPPELINSTINIHRTSVVEGLCEGIGA 240
QY      241 GLVDVAIWGTGSDYPRKGAISTGMNSVSRITIEELPK 277
      |||
Db       241 GLVDVAIWGTGSDYPRKGAISTGMNSVSRITIEELPK 277

RESULT 2
US-09-692-081-2
; Sequence 2, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
```

```
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; CURRENT FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-692-081-2
```

```
Query Match      88.6%; Score 1307; DB 4; Length 245;
Best Local Similarity 100.0%; Pred. No. 1.8e-132;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      33 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVKQKALIRREVVDLYNGMCLQG 92
DB      1 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVKQKALIRREVVDLYNGMCLQG 60
QY      93 PAGVPGRDSPGANGIPGTGPIGRDGFKEGECLEESFEESWTPTYKQCSWSSLYNGI 152
DB      61 PAGVPGRDSPGANGIPGTGPIGRDGFKEGECLEESFEESWTPTYKQCSWSSLYNGI 120
QY      153 DGIKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWFTFNAGCSGPLPIEAIIVL 212
DB      121 DGIKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWFTFNAGCSGPLPIEAIIVL 180
QY      213 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRIT 272
DB      181 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRIT 240
QY      273 EELPK 277
DB      241 EELPK 245
```

```
RESULT 3
US-09-834-759-515
; Sequence 515, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
```

```
; APPLICANT: JIANG, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-515
```

```
Query Match      83.6%; Score 1234.5; DB 4; Length 278;
Best Local Similarity 86.4%; Pred. No. 1.4e-124;
Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;
```

```
QY      1 MRPAEL---GQLTSRAGLCRPLCLLCASQLPHTMHPORASPOLLLGLFVLLLLQ 57
DB      1 MDPAAASRGGAADHVPLIGLRLQURAKRQGANRPOGPASPOLRLG--LILLILLQ 58
QY      58 LSA PSSASENPKVKQKALIRREVVDLYNGMCLQGAGVPGRDSPGANGIPGTGIPGR 117
```

```
DB      59 LPAPSSASEIPKQKQKQALQREVVDLYNGMCLQGPAGVPGRDSPGANVIRPGTGPGR 118
QY      118 DGFKEGECLEESFEESWTPTYKQCSWSSLYNGIDLGKIAECTFTKRSNSALRVLSG 177
DB      119 DGFKEGECLEESFEESWTPTYKQCSWSSLYNGIDLGKIAECTFTKRSNSALRVLSG 178
QY      178 SLRLKCRNACCORWFTFNAGCSGPLPIEAIIVDOGSPELNSTINIHRTSSVEGLCEG 237
DB      179 SLRLKCRNACCORWFTFNAGCSGPLPIEAIIVDOGSPELNSTINIHRTSSVEGLCEG 238
QY      238 IGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITIEELPK 277
DB      239 IGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRITIEELPK 278
```

```
RESULT 4
US-09-692-081-4
; Sequence 4, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
```

```
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; CURRENT FILING DATE: 2000-10-19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-692-081-4
```

```
Query Match      83.0%; Score 1225; DB 4; Length 243;
Best Local Similarity 94.7%; Pred. No. 1.2e-123;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```
QY      33 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVKQKALIRREVVDLYNGMCLQG 92
DB      1 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASEIPKQKQKQALQREVVDLYNGMCLQG 58
QY      93 PAGVPGRDSPGANGIPGTGPIGRDGFKEGECLEESFEESWTPTYKQCSWSSLYNGI 152
DB      59 PAGVPGRDSPGANGIPGTGPIGRDGFKEGECLEESFEESWTPTYKQCSWSSLYNGI 118
QY      153 DGIKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWFTFNAGCSGPLPIEAIIVL 212
DB      119 DGIKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWFTFNAGCSGPLPIEAIIVL 178
QY      213 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRIT 272
DB      179 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRIT 238
QY      273 EELPK 277
DB      239 EELPK 243
```

```
RESULT 5
US-09-489-847-205
; Sequence 205, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
```

```
; APPLICANT: ROSEN et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
```

```

; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 205
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-205
```

```

Query Match      82.4%; Score 1222; DB 4; Length 243;
Best Local Similarity 94.3%; Pred. No. 2.6e-123;
Matches 231; Conservative 4; Mismatches 8; Indels 2; Gaps 1;
```

```

QY 33 MHPOGSAASPOLLLGLFVLVLLLLQLSABSSASBNPKVKKALIRQREVVLDYNGMCTOG 92
DB 1 MRPOGPAASPORLGG--LILLLLLLQLPAPSSASEIPKQKQKQLRQREVVLDYNGMCTOG 58
QY 93 PAVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESFEESWTPNYKCCSSSLNYGI 152
DB 59 PAVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESFEESWTPNYKCCSSSLNYGI 118
QY 153 DLGKIAECTFTTKRNSALRVLFSSGLRLKCRNACCORWYFTTNGACSGPLPIEAIYYL 212
DB 119 DLGKIAECTFTTKRNSALRVLFSSGLRLKCRNACCORWYFTTNGACSGPLPIEAIYYL 178
QY 213 DQSPPLNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIT 272
DB 179 DQSPPLNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIT 238
QY 273 EELPK 277
DB 239 EELPK 243
```

```

RESULT 6
US-09-834-759-514
; Sequence 514, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121, 470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 514
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-514
```

```

Query Match      82.4%; Score 1216; DB 4; Length 243;
Best Local Similarity 94.3%; Pred. No. 1.1e-122;
Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1;
```

```

QY 33 MHPOGSAASPOLLLGLFVLVLLLLQLSABSSASBNPKVKKALIRQREVVLDYNGMCTOG 92
DB 1 MRPOGPAASPORLGG--LILLLLLLQLPAPSSASEIPKQKQKQLRQREVVLDYNGMCTOG 58
```

```

QY 93 PAVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESFEESWTPNYKCCSSSLNYGI 152
DB 59 PAVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESFEESWTPNYKCCSSSLNYGI 118
QY 153 DLGKIAECTFTTKRNSALRVLFSSGLRLKCRNACCORWYFTTNGACSGPLPIEAIYYL 212
DB 119 DLGKIAECTFTTKRNSALRVLFSSGLRLKCRNACCORWYFTTNGACSGPLPIEAIYYL 178
QY 213 DQSPPLNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIT 272
DB 179 DQSPPLNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIT 238
QY 273 EELPK 277
DB 239 EELPK 243
```

```

RESULT 7
US-09-205-258-958
; Sequence 958, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young, et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 958
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-938

```

```

Query Match          74.8%; Score 1103.5; DB 4; Length 276;
Best Local Similarity 74.9%; Pred. No. 1.8e-110;
Matches 221; Conservative 6; Mismatches 27; Indels 41; Gaps 5;

```

```

QY 2 RPAELGQTL-----SRAGLCRPLCLLCASQDPTHMPOGRAASPOLLLGLFLVLLI 54
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4 RPSAAXQTLTTLPLSSVSSAS-----SSALPGSRPPCDPPRAPPPRSG----- 46
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 55 LQLLSAPSSAS-----ENPKVKQKALIRQEVVDLYNGMCLGSPAGVGRDGS 102
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 47 ----SAAACSCSCSCCPRRRAPLRSPR-GSKRIKQREVVLDLNGMCLGSPAGVGRDGS 101
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 103 PGANGIPGTPIGPDGFKGKEKCECLRESFEESWTNPKOCSSWSLNYGIDLKIAECTF 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 102 PGANGIPGTPIGPDGFKGKEKCECLRESFEESWTNPKOCSSWSLNYGIDLKIAECTF 161
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 163 TKRNSALRYLFFSGSLRLKCRACCCORWFTFNGABCSGPLEIAIITYDQSPPELNST 222
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 162 TKRNSALRYLFFSGSLRLKCRACCCORWFTFNGABCSGPLEIAIITYDQSPPELNST 221
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 223 INHRTSSVYGLCEGIGAGLVDAIWMGTCSDPKGDASTGNSVSRITIEELPK 277
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 222 INHRTSSVYGLCEGIGAGLVDAIWMGTCSDPKGDASTGNSVSRITIEELPK 276
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 8

```

US-09-834-759-517
; Sequence 517, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepier, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 517
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-517

```

```

Query Match          66.8%; Score 986.5; DB 4; Length 232;
Best Local Similarity 83.8%; Pred. No. 5.7e-98;
Matches 196; Conservative 6; Mismatches 27; Indels 5; Gaps 2;

```

```

QY 1 MRPAEL--GOTLSRAGLCRPLCLLCASQDPTHMPOGRAASPOLLLGLFLVLLI 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MQPAASERGADADHVPLGLRLQLRAQPGAMRPGAPASQRLRG--LTLILLIQ 58
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 58 LSAPSSASNPVKQKALIRQEVVDLYNGMCLGSPAGVGRDGS PGANGIPGTGIPGR 117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 59 LPAPSSASRIPKQKALIRQEVVDLYNGMCLGSPAGVGRDGS PGANVLPGTGIPGR 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 DGFKEKGECLRESFEESWTNPKOCSSWSLNYGIDLKIAECTFTKRNSALRYLPSG 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 119 DGFKEKGECLRESFEESWTNPKOCSSWSLNYGIDLKIAECTFTKRNSALRYLPSG 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 SRLKCRNACCORWFTFNGABCSGPLEIAIITYDQSPPELNSTINHRTSSV 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 179 SRLKCRNACCORWFTFNGABCSGPLEIAIITYDQSPPELNSTINHRTSSV 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 9
US-09-834-759-516
; Sequence 516, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepier, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 516
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-516

```

```

Query Match          65.6%; Score 968; DB 4; Length 197;
Best Local Similarity 93.0%; Pred. No. 4.4e-96;
Matches 185; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

```

```
QY 33 MHPOGRASPOLLLGLFLVLLLLQLSAPSSASBNPKVKOKALIROREVDLYNMGKLOG 92
DB 1 MRQGRASPOLRLG--LTLALLQLDPASSASIPGKKQKQALROREVDLYNMGKLOG 58
QY 93 PAGVPGRDSPGANVIGPTGPIGRDGFKEGKECLRESFEESWTPNYKCSMSLNYGI 152
DB 59 PAGVPGRDSPGANVIGPTGPIGRDGFKEGKECLRESFEESWTPNYKCSMSLNYGI 118
QY 153 DLKIAECTTKRNSALNVLFGSGLRKCRRNACCORWFTTNGAECGPLEIAIYL 212
DB 119 DLKIAECTTKRNSALNVLFGSGLRKCRRNACCORWFTTNGAECGPLEIAIYL 178
QY 213 DOGSEPLNSTINIRTSV 231
DB 179 DOGSEPLNSTINIRTSV 197

RESULT 10
US-09-205-258-962
; Sequence 962, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,557
; EARLIER FILING DATE: 1998-07-30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 962
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-962

Query Match 23.9%; Score 353; DB 4; Length 66;
Best Local Similarity 98.5%; Pred. No. 1.7e-30;
Matches 65; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 162 FTKRMSNSALRYVLFSGSLRLKCRNACCORWFTTNGAECGPLEIAIYLLDOGSEPLNS 221
DB 1 FTKRMSNSALRYVLFSGSLRLKCRNACCORWFTTNGAECGPLEIAIYLLDOGSEPLNS 60
QY 222 TINIR 227
DB 61 TINIR 66

RESULT 11
US-09-205-258-961
; Sequence 961, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
```

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30

NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 961  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-205-258-961  
  
Query Match 20.1%; Score 296; DB 4; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1,7e-24;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 111 TPGI PRGDFGKGEKCECLRESFESBSMTNPKQCSWSLNYGIDGKIACTF 162  
Db 1 TPGI PRGDFGKGEKCECLRESFESBSMTNPKQCSWSLNYGIDGKIACTF 52  
  
RESULT 12  
US-09-205-258-963  
Sequence 963, Application US/09205258  
Patent No. 6525174  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P2007P1  
CURRENT APPLICATION NUMBER: US/09/205,258  
CURRENT FILING DATE: 1998-12-04  
EARLIER APPLICATION NUMBER: FCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06

```

; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 963
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-205-258-963

Query Match      18.3%  Score 270; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 1e-21;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      227 RTSSVGLCEGIGAGLVDAIHWGTCSDYKPGDASTGMSVSRIIIIEELPK 277
DB      1 RTSSVGLCEGIGAGLVDAIHWGTCSDYKPGDASTGMSVSRIIIIEELPK 51

RESULT 13
; US-09-834-759-518
; Sequence 518, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 518
; LENGTH: 46
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
; US-09-834-759-518

Query Match      16.8%  Score 248; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      232 EGICEGIGAGLVDAIHWGTCSDYKPGDASTGMSVSRIIIIEELPK 277
DB      1 EGICEGIGAGLVDAIHWGTCSDYKPGDASTGMSVSRIIIIEELPK 46

RESULT 14
; US-09-205-258-960
; Sequence 960, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
```

```

: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/049,374
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,917
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,949
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,974
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,883
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,897
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,898
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,962
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,963
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,877
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/048,878
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/070,923
: EARLIER FILING DATE: 1997-12-18
: EARLIER APPLICATION NUMBER: 60/092,921
: EARLIER FILING DATE: 1998-07-15
: EARLIER APPLICATION NUMBER: 60/094,657
: EARLIER FILING DATE: 1998-07-30
: NUMBER OF SEQ ID NOS: 1227
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 960
: LENGTH: 52
: TYPE: prt
: ORGANISM: Homo sapiens
: US-09-205-258-960

```

	Query Match	Similarity	Score	DB	Length
Best Local	40	95.2%	Pred	No. 8.9e-16	52
Matches	40	Conservative	0	Mismatches	2
				Indels	0
				Gaps	0
Oy	73	KALIRQREYVDLYNGMCILOGPAGVDPGRDQSPGANGSI	PGTPTGI	114	
Ob	11	KRIRIQREYVDLYNGMCILOGPAGVDPGRDQSPGANGSI	PGTPTGI	52	

```

RESULT 15
US-09-834-759-519
; Sequence 519, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.47009
; CURRENT APPLICATION NUMBER: US/09/834, 759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 519
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-519

Query Match          9.6%; Score 141; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 3e-08;

```

	Matches	26,	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
QY	252	CSDYPKGDASTGMSVSRITIEEL.PK	277							
Dd	1	CSDYPKGDASTGMSVSRITIEEL.PK	26							

Search completed: August 4, 2005, 04:23:54  
Job time : 28.881 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:21:44 ; Search time 98.1268 Seconds  
(without alignments)  
1099.816 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476

Sequence: 1 MRPAELGQTLSRAGLCRPL.....GDASTGMSVSRITIELPK 277

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1745140 seqs, 389608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubppa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubppa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubppa/US10E\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/2/pubppa/US10F\_PUBCOMB.pep:\*  
19: /cgn2\_6/ptodata/2/pubppa/US11\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/2/pubppa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/2/pubppa/US11\_PUBCOMB.pep:\*  
22: /cgn2\_6/ptodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1476	100.0	277	US-10-045-992-5	Sequence 5, Appli
2	1476	100.0	277	US-10-634-108-5	Sequence 5, Appli
3	1476	100.0	277	US-10-939-233-5	Sequence 5, Appli
4	1307	88.6	245	US-10-045-992-2	Sequence 2, Appli
5	1307	88.6	245	US-10-634-108-2	Sequence 2, Appli
6	1307	88.6	245	US-10-939-233-2	Sequence 2, Appli
7	1243.5	84.2	278	US-10-097-340-131	Sequence 131, App
8	1243.5	84.2	278	US-10-177-293-496	Sequence 496, App
9	1243.5	84.2	278	US-10-301-822-77	Sequence 77, Appli
10	1243.5	84.2	278	US-10-961-139-2	Sequence 2, Appli
11	1240.5	84.0	278	US-10-060-036-4551	Sequence 4551, Ap

12	1239.5	84.0	278	15	US-10-296-115-1261	Sequence 1261, Ap
13	1234.5	83.6	278	9	US-09-834-759-515	Sequence 515, App
14	1234.5	83.6	278	13	US-10-007-805-515	Sequence 515, App
15	1234.5	83.6	278	14	US-10-076-622-515	Sequence 515, App
16	1234.5	83.6	278	14	US-10-060-036-4560	Sequence 4560, Ap
17	1234.5	83.6	278	14	US-10-124-805-515	Sequence 515, App
18	1225	83.0	243	12	US-09-829-472A-19	Sequence 19, Appli
19	1225	83.0	243	13	US-10-045-992-4	Sequence 4, Appli
20	1225	83.0	243	14	US-10-097-340-135	Sequence 135, App
21	1225	83.0	243	15	US-10-295-027-50	Sequence 50, Appli
22	1225	83.0	243	15	US-10-173-999-74	Sequence 74, Appli
23	1225	83.0	243	15	US-10-058-270A-38	Sequence 38, Appli
24	1225	83.0	243	15	US-10-634-108-4	Sequence 4, Appli
25	1225	83.0	243	15	US-10-188-832-115	Sequence 115, App
26	1225	83.0	243	15	US-10-176-758-366	Sequence 366, App
27	1225	83.0	243	18	US-10-351-334-205	Sequence 205, App
28	1222	82.8	243	15	US-10-176-847-14	Sequence 14, Appli
29	1220	82.7	243	14	US-09-834-759-514	Sequence 514, App
30	1216	82.4	243	9	US-09-938-418-7	Sequence 7, Appli
31	1216	82.4	243	10	US-09-946-374-352	Sequence 352, App
32	1216	82.4	243	10	US-09-946-374-431	Sequence 431, App
33	1216	82.4	243	13	US-10-006-867-122	Sequence 122, App
34	1216	82.4	243	13	US-10-052-586-366	Sequence 366, App
35	1216	82.4	243	13	US-10-007-805-514	Sequence 514, App
36	1216	82.4	243	13	US-10-063-547-122	Sequence 122, App
37	1216	82.4	243	13	US-10-063-551-122	Sequence 122, App
38	1216	82.4	243	14	US-10-174-590-366	Sequence 366, App
39	1216	82.4	243	14	US-10-175-737-366	Sequence 366, App
40	1216	82.4	243	14	US-10-176-758-366	Sequence 366, App
41	1216	82.4	243	14	US-10-063-616-122	Sequence 122, App
42	1216	82.4	243	14	US-10-174-581-366	Sequence 366, App
43	1216	82.4	243	14	US-10-176-483-366	Sequence 366, App
44	1216	82.4	243	14	US-10-176-749-366	Sequence 366, App
45	1216	82.4	243	14	US-10-176-749-366	Sequence 366, App

#### ALIGNMENTS

```
RESULT 1
US-10-045-992-5
; Sequence 5, Application US//10045992
; Publication No. US20020161211A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: PRISEU, Robert E.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN
; FILE REFERENCE: 053689-5006-01
; CURRENT APPLICATION NUMBER: US/10/045,992
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-045-992-5

Query Match      100.0%; Score 1476; DB 13; Length 277;
Best Local Similarity 100.0%; Pred. No. 3.4e-134;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPAELGQTLSRAGLCRPLCLLCASQLPHMHPQGRASPQLLGLFLVLLLLQSLA 60
        |||||
DB      1 MRPAELGQTLSRAGLCRPLCLLCASQLPHMHPQGRASPQLLGLFLVLLLLQSLA 60
        |||||

QY      61 PSSASBNPKVKKALIRQREVVLDYNGCLQCPAGVPGRDSFGANGIPGTGIGRDRF 120
        |||||
DB      61 PSSASBNPKVKKALIRQREVVLDYNGCLQCPAGVPGRDSFGANGIPGTGIGRDRF 120
        |||||

QY      121 KKKKCKLRSPSESWTPVYKQCSMSLNYGIDLKIAECFTTKRNSALRVLPSGSLR 180
        |||||
```

Db 121 KGEKECLRESFEESWTPTNYKQCSWSLNYGIDLGKIACTFTKMSNSALRVLFSSGLR 180  
Qy 181 LKCRNACCORWYFTTNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEGIGA 240  
Db 181 LKCRNACCORWYFTTNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEGIGA 240  
Qy 241 GLVDVAIWGTCSDYKPGDASTGMSNSVRITIIIEELPK 277  
Db 241 GLVDVAIWGTCSDYKPGDASTGMSNSVRITIIIEELPK 277

## RESULT 2

US-10-634-108-5  
; Sequence 5, Application US/10634108  
; Publication No. US20040063176A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volkhart  
; APPLICANT: FRIESEL, Robert F.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL  
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)  
; CURRENT APPLICATION NUMBER: US/10/634,108  
; PRIOR FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/692,081  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-634-108-5

Query Match 100.0%; Score 1476; DB 15; Length 277;  
Best Local Similarity 100.0%; Pred. No. 3.4e-134;  
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPPAEELGOTLSRAGLCRPLCLILCASOLPHTMHPQGRASPOLLGLFLVLLILLQLQSA 60  
Db 1 MPPAEELGOTLSRAGLCRPLCLILCASOLPHTMHPQGRASPOLLGLFLVLLILLQLQSA 60  
Qy 61 PSSASENPVKOKALIRQREVVDLYNGMCLQGPAGVPGRDSPGANGIPGTPIGRDGF 120  
Db 61 PSSASENPVKOKALIRQREVVDLYNGMCLQGPAGVPGRDSPGANGIPGTPIGRDGF 120  
Qy 121 KGEKECLRESFEESWTPTNYKQCSWSLNYGIDLGKIACTFTKMSNSALRVLFSSGLR 180  
Db 121 KGEKECLRESFEESWTPTNYKQCSWSLNYGIDLGKIACTFTKMSNSALRVLFSSGLR 180  
Qy 181 LKCRNACCORWYFTTNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEGIGA 240  
Db 181 LKCRNACCORWYFTTNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEGIGA 240  
Qy 241 GLVDVAIWGTCSDYKPGDASTGMSNSVRITIIIEELPK 277  
Db 241 GLVDVAIWGTCSDYKPGDASTGMSNSVRITIIIEELPK 277

RESULT 3  
US-10-939-233-5  
; Sequence 5, Application US/10939233  
; Publication No. US20050147602A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volkhart  
; APPLICANT: FRIESEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTNRC1, A NOVEL  
; FILE REFERENCE: 053689-5006-03  
; CURRENT APPLICATION NUMBER: US/10/939,233  
; PRIOR FILING DATE: 2004-09-10  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US 10/045,992

; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-939-233-5

Query Match 100.0%; Score 1476; DB 18; Length 277;  
Best Local Similarity 100.0%; Pred. No. 3.4e-134;  
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPPAEELGOTLSRAGLCRPLCLILCASOLPHTMHPQGRASPOLLGLFLVLLILLQLQSA 60  
Db 1 MPPAEELGOTLSRAGLCRPLCLILCASOLPHTMHPQGRASPOLLGLFLVLLILLQLQSA 60  
Qy 61 PSSASENPVKOKALIRQREVVDLYNGMCLQGPAGVPGRDSPGANGIPGTPIGRDGF 120  
Db 61 PSSASENPVKOKALIRQREVVDLYNGMCLQGPAGVPGRDSPGANGIPGTPIGRDGF 120  
Qy 121 KGEKECLRESFEESWTPTNYKQCSWSLNYGIDLGKIACTFTKMSNSALRVLFSSGLR 180  
Db 121 KGEKECLRESFEESWTPTNYKQCSWSLNYGIDLGKIACTFTKMSNSALRVLFSSGLR 180  
Qy 181 LKCRNACCORWYFTTNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEGIGA 240  
Db 181 LKCRNACCORWYFTTNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEGIGA 240  
Qy 241 GLVDVAIWGTCSDYKPGDASTGMSNSVRITIIIEELPK 277  
Db 241 GLVDVAIWGTCSDYKPGDASTGMSNSVRITIIIEELPK 277

## RESULT 4

US-10-045-992-2  
; Sequence 2, Application US/10045992  
; Publication No. US20020161211A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volkhart  
; APPLICANT: FRIESEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN  
; FILE REFERENCE: 053689-5006-01  
; CURRENT APPLICATION NUMBER: US/10/045,992  
; PRIOR FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-045-992-2

Query Match 88.6%; Score 1307; DB 13; Length 245;  
Best Local Similarity 100.0%; Pred. No. 6.7e-118;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 MHPQGRASPOLLGLFLVLLILLQLSAPSSASENPVKOKALIRQREVVDLYNGMCLQ 92  
Db 1 MHPQGRASPOLLGLFLVLLILLQLSAPSSASENPVKOKALIRQREVVDLYNGMCLQ 60  
Qy 93 PAVGVRGDSPPANGIPGTPIGRDGRKGEKGLRESFEESWTPTNYKQCSWSLNYGI 152  
Db 61 PAVGVRGDSPPANGIPGTPIGRDGRKGEKGLRESFEESWTPTNYKQCSWSLNYGI 120  
Qy 153 DKGIAECTFTKMSNSALRVLFSSGLRLKCRNACCORWYFTTNGAECGSLPIEALIIYL 212  
Db 121 DKGIAECTFTKMSNSALRVLFSSGLRLKCRNACCORWYFTTNGAECGSLPIEALIIYL 180  
Qy 213 DQSPPELNTSTINIHRTSSVEGLCEGIGLVDVAIWGTCSDYKPGDASTGMSNSVRITII 272

Db 181 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSR111 240  
QY 273 EELPK 277  
Db 241 EELPK 245

RESULT 5  
US-10-634-108-2

; Sequence 2, Application US/10634108  
; Publication No. US20040063176A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volkhard  
; APPLICANT: PRIESEL, Robert P.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL  
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)  
; CURRENT APPLICATION NUMBER: US/10/634,108  
; PRIOR FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/692,081  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-634-108-2

Query Match 88.6%; Score 1307; DB 15; Length 245;  
Best Local Similarity 100.0%; Pred. No. 6,7e-118;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MHPOGRASPOLLLGLFLVLLLLQLSAPSSASENPVKOKALIRORVVDLYNMCLOG 92  
Db 1 MHPOGRASPOLLLGLFLVLLLLQLSAPSSASENPVKOKALIRORVVDLYNMCLOG 60  
QY 93 PAVPGRDGPNGANGIPGTGPIGRDGFKEGKECELRSESESWTPYKCCSSSLNYGI 152  
Db 61 PAVPGRDGPNGANGIPGTGPIGRDGFKEGKECELRSESESWTPYKCCSSSLNYGI 120  
QY 153 DLGKIACEFTTKRNSALRVLFSSGLRLKCRNACCORWFTTNGACSGPLIEATIIYL 212  
Db 121 DLGKIACEFTTKRNSALRVLFSSGLRLKCRNACCORWFTTNGACSGPLIEATIIYL 180  
QY 213 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSR111 272  
Db 181 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSR111 240  
QY 273 EELPK 277  
Db 241 EELPK 245

RESULT 6  
US-10-939-233-2  
; Sequence 2, Application US/10939233  
; Publication No. US20050147602A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volkhard  
; APPLICANT: PRIESEL, Robert P.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTNRC1, A NOVEL  
; TITLE OF INVENTION: MODULATOR OF COLLAGEN MATRIX  
; FILE REFERENCE: 053689-5006-03  
; CURRENT APPLICATION NUMBER: US/10/939,233  
; CURRENT FILING DATE: 2004-09-10  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US 10/045,992  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 245

; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-939-233-2

Query Match 88.6%; Score 1307; DB 18; Length 245;  
Best Local Similarity 100.0%; Pred. No. 6,7e-118;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MHPOGRASPOLLLGLFLVLLLLQLSAPSSASENPVKOKALIRORVVDLYNMCLOG 92  
Db 1 MHPOGRASPOLLLGLFLVLLLLQLSAPSSASENPVKOKALIRORVVDLYNMCLOG 60  
QY 93 PAVPGRDGPNGANGIPGTGPIGRDGFKEGKECELRSESESWTPYKCCSSSLNYGI 152  
Db 61 PAVPGRDGPNGANGIPGTGPIGRDGFKEGKECELRSESESWTPYKCCSSSLNYGI 120  
QY 153 DLGKIACEFTTKRNSALRVLFSSGLRLKCRNACCORWFTTNGACSGPLIEATIIYL 212  
Db 121 DLGKIACEFTTKRNSALRVLFSSGLRLKCRNACCORWFTTNGACSGPLIEATIIYL 180  
QY 213 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSR111 272  
Db 181 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPKGDASTGMNSVSR111 240  
QY 273 EELPK 277  
Db 241 EELPK 245

RESULT 7  
US-10-097-340-131

; Sequence 131, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVAPU  
; APPLICANT: Sebastian HOERSCHE  
; APPLICANT: Shubhangi KAMATKAR  
; APPLICANT: Steve G. KOVATIS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VERIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LV  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
; FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; CURRENT FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/325,102  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/323,580  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 131  
; LENGTH: 278

	TYPE: PRT								
		ORGANISM: Homo sapiens							
		US-10-097-340-131							
Query Match:	84.2%	Score 1243.5;	DB 14;	Length 278;					
Best Local Similarity	86.8%	Pred. No. 1,1e-111;							
Matches 243;	Conservative	6;	Mismatches 26;	Indels 5;	Gaps 2;				
QY	1	MRPAAL---GQTLSRAGLCRPLCLLLCASQLPHMTBQGRASPOLLGLGFLVLLDLO	57						
DB	1	MQPAASBRGGADADHVPLGLRLQLRAARQPMGRQGPASPQRURG--LILLILLDQ	58						
QY	58	LSAPSSAENPRVYKQKALIRQREVDVLYNGMQLQSPAVPGRDGPANGIPGTGIGCR	117						
DB	59	LPAPSSAEIPKQKQKALIRQREVDVLYNGMQLQSPAVPGRDGPANGIPGTGIGCR	118						
QY	118	DGFKEKEGCLRESPEESMTPIRYKOCSSSLNYNGIDGLKIAECTFTTKRSNSALRVLPFG	177						
DB	119	DGFKEKEGCLRESPEESMTPIRYKOCSSSLNYNGIDGLKIAECTFTTKRSNSALRVLPFG	178						
QY	178	SLRLKRNACCORWYFTFNGAECSGPLIEALIVLDQSPBELNSTITINIRHTSSVEGLCEG	237						
DB	179	SLRLKRNACCORWYFTFNGAECSGPLIEALIVLDQSPBELNSTITINIRHTSSVEGLCEG	238						
QY	238	IGAGLVDAIVAVTGCTSDYPKGDASTGNNVSRIIIEELPK	277						
DB	239	IGAGLVDAIVAVTGCTSDYPKGDASTGNNVSRIIIEELPK	278						

```

      RESULT 8
      US-10-177-293-496
      ; Sequence 496, Application US/10177293
      ; Publication No. US20030124128A1
      ; GENERAL INFORMATION:
      ; APPLICANT: Lillie, James
      ; APPLICANT: Glatz, Karen
      ; APPLICANT: Zhao, Xumei
      ; APPLICANT: Ganavarpur, Manjula
      ; APPLICANT: Kamatekar, Shubhangit
      ; APPLICANT: Mertens, Maureen
      ; APPLICANT: Myer, Vic
      ; APPLICANT: Wang, Youzhen
      ; APPLICANT: Xu, Yongyao
      ; APPLICANT: Hoersch, Sebastian
      ; APPLICANT: Monahan, John
      ; APPLICANT: Meyers, Rachel E.
      ; APPLICANT: Bast Jr., Robert C.
      ; APPLICANT: Hortobagyi, Gabriel N.
      ; APPLICANT: Pusztai, Lajos
      ; APPLICANT: Meric, Funda
      ; APPLICANT: Sahin, Aysegul
      ; APPLICANT: Mills, Gordon B.
      ; TITLE OF INVENTION: PREVENTION, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
      ; FILE REFERENCE: MRI-038
      ; CURRENT APPLICATION NUMBER: US/10/177,293
      ; CURRENT FILING DATE: 2002-06-21
      ; PRIOR APPLICATION NUMBER: US 60/299,887
      ; PRIOR FILING DATE: 2001-06-21
      ; PRIOR APPLICATION NUMBER: US 60/301,572
      ; PRIOR FILING DATE: 2001-06-27
      ; PRIOR APPLICATION NUMBER: US 60/306,501
      ; PRIOR FILING DATE: 2001-07-18
      ; PRIOR APPLICATION NUMBER: US 60/325,002
      ; PRIOR FILING DATE: 2001-09-25
      ; PRIOR APPLICATION NUMBER: US 60/362,585
      ; PRIOR FILING DATE: 2002-03-05
      ; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
      ; PRIOR FILING DATE: 2002-05-14
      ; NUMBER OF SEQ ID NOS: 506
      ; SOFTWARE: FaetSeq for Windows Version 4.0
      ; SEQ ID NO 496
      ; LENGTH: 278

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-496

Query Match      84.2%; Score 1243.5; DB 14; Length 278;
Best local similarity 86.8%; Pred. No. 1,1e-111;
Matches 243; Conservative 6; Mismatches 26; Indels 5; Gaps 2;

QY      1 MRPAET---GQTSRAGLCRPLCLLLCASQLPHTMHPOGPAASPQLLGLFVLLLLQ 57
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 MQPAASERGGADHDHVDPLGLTLRLQLTQRAAQPGAMFPQGPAAAPQRFARG--LILLLLDQ 58

QY      58 LSAPSSASENKKVKOKLILROREVDVLYNGCICLGGPGVPERDSSPGANGIPGTGPR 117
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      59 LPAPSASEIPKQKOKQLROREVDVLYNGCICLGGPGVPERDSSPGANGIPGTGPR 118

QY      118 DGFKEKGECLRESFEESGWTPNYVQCQSSSLNTYGDILGKIAECTFTTKRSNSALRVLPFG 177
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      119 DGFKEKGECLRESFEESGWTPNYVQCQSSSLNTYGDILGKIAECTFTTKRSNSALRVLPFG 178

QY      178 SLRLKRNACCQRMVFTFNGAECGGLPIEAILIYLDQSPFLNSTINIHRTSSVBLCEG 237
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      179 SLRLKRNACCQRMVFTFNGAECGGLPIEAILIYLDQSPFLNSTINIHRTSSVBLCEG 238

QY      238 IGAGLVDAIVAVGTCSDYPKGDASTGMNSVRIIIEELPK 277
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      239 IGAGLVDAIVAVGTCSDYPKGDASTGMNSVRIIIEELPK 278

```

[illegible]



```
Db 21 GLLR---LQLRAARQPGAMRPGOPAASPOLRLG--LTLTLTLQLPAPSSASEIPKQKQA 75
Qy 75 LIRQREVVDLYNMCLOQAGVPGRGSGANGIPCTPGIPGRDGKGEKGLRSEFE 134
Db 76 QLRQREVVDLYNMCLOQAGVPGRGSGANGIPCTPGIPGRDGKGEKGLRSEFE 135
Qy 135 SWTPNKKOCSSWSLNYGIDLGKIAECTPTKMRNSALRYLFSGSLTKCRNACCOMWYFT 194
Db 136 SWTPNKKOCSSWSLNYGIDLGKIAECTPTKMRNSALRYLFSGSLTKCRNACCOMWYFT 195
Qy 195 FNGAECGSLPIEAIYYLDQGSPELNTINIHRTSSVEGLCEGIGALVDVAIWGTCSG 254
Db 196 FNGAECGSLPIEAIYYLDQGSPELNTINIHRTSSVEGLCEGIGALVDVAIWGTCSG 255
Qy 255 YPKGASTGWSVSRITIEELPK 277
Db 256 YPKGASTGWSVSRITIEELPK 278

RESULT 13
US-09-834-759-515
; Sequence 515, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834, 759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-515

Query Match 83.6%; Score 1234.5; DB 9; Length 278;
Best Local Similarity 86.4%; Pred. No. 8.1e-111;
Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;

Qy 1 MRPAEL---GOTLSRAGLCRPLCLLCA SOLPHTMHPOGRAASPOLLIGLFLVLLLLIQ 57
Db 1 MQPAAASERGAADADHVPILGLRLQLRAARQPGAMRPGOPAASPOLRLG--LTLTLTLQ 58
Qy 58 LSAPSSASENPKYKQKALIRQREVVDLYNMCLOQAGVPGRGSGANGIPCTPGIPGR 117
Db 59 LPAPSSASEIPKQKQKALRQREVVDLYNMCLOQAGVPGRGSGANGIPCTPGIPGR 118
Qy 118 DGFKEGKECLRESFEESWTPNKKOCSSWSLNYGIDLGKIAECTPTKMRNSALRYLFSG 177
Db 119 DGFKEGKECLRESFEESWTPNKKOCSSWSLNYGIDLGKIAECTPTKMRNSALRYLFSG 178
Qy 178 SLRLKCRNACCOMWYFTFNGAECGSLPIEAIYYLDQGSPELNTINIHRTSSVEGLCEG 237
Db 179 SLRLKCRNACCOMWYFTFNGAECGSLPIEAIYYLDQGSPELNTINIHRTSSVEGLCEG 238
Qy 238 IGAGLVDAVAIWGTCSGDIYPRKGDASTGMNSVSRITIEELPK 277
Db 239 IGAGLVDAVAIWGTCSGDIYPRKGDASTGMNSVSRITIEELPK 278

RESULT 14
US-10-007-805-515
; Sequence 515, Application US/10007805
; Publication No. US20020150581A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margareta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-515

Query Match 83.6%; Score 1234.5; DB 13; Length 278;
Best Local Similarity 86.4%; Pred. No. 8.1e-111;
Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;

Qy 1 MRPAEL---GOTLSRAGLCRPLCLLCA SOLPHTMHPOGRAASPOLLIGLFLVLLLLIQ 57
Db 1 MQPAAASERGAADADHVPILGLRLQLRAARQPGAMRPGOPAASPOLRLG--LTLTLTLQ 58
Qy 58 LSAPSSASENPKYKQKALIRQREVVDLYNMCLOQAGVPGRGSGANGIPCTPGIPGR 117
Db 59 LPAPSSASEIPKQKQKALRQREVVDLYNMCLOQAGVPGRGSGANGIPCTPGIPGR 118
Qy 118 DGFKEGKECLRESFEESWTPNKKOCSSWSLNYGIDLGKIAECTPTKMRNSALRYLFSG 177
Db 119 DGFKEGKECLRESFEESWTPNKKOCSSWSLNYGIDLGKIAECTPTKMRNSALRYLFSG 178
Qy 178 SLRLKCRNACCOMWYFTFNGAECGSLPIEAIYYLDQGSPELNTINIHRTSSVEGLCEG 237
Db 179 SLRLKCRNACCOMWYFTFNGAECGSLPIEAIYYLDQGSPELNTINIHRTSSVEGLCEG 238
Qy 238 IGAGLVDAVAIWGTCSGDIYPRKGDASTGMNSVSRITIEELPK 277
Db 239 IGAGLVDAVAIWGTCSGDIYPRKGDASTGMNSVSRITIEELPK 278

RESULT 15
US-10-076-622-515
; Sequence 515, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Persing, David H.
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-515

Query Match 83.6%; Score 1234.5; DB 14; Length 278;
Best Local Similarity 86.4%; Pred. No. 8.1e-111;
```

Matches	242;	Conservative	6;	Mismatches	27;	Indels	5;	Gaps	2;
Qy	1	MRPAAL---	QOTLSRAGLCRPLCLLLCASQSPHTMHQGRASPQLLSGLFVYLILLQ						57
Db	1	MOPAASERGDADHVPILGLRLQLPAAKPGAMRQGPASQRRRG---	LLLLLLLQ						58
Qy	58	LSAPSSASENKVKOKALIRQREVVDLYNGMCLQGPAGVPGRDGSPGANGIPGTGPIGR							117
Db	59	LPASSASEIPKQKQKQALQREVVDLYNGMCLQGPAGVPGRDGSPGANVIPGTGPIGR							118
Qy	118	DGFGEKEKCELRSESEBSWTPNYKOCSSSNLYNGIDLKIAECTPTTKRSSALRVLPFG							177
Db	119	DGFGEKEKCELRSESEBSWTPNYKOCSSSNLYNGIDLKIAECTPTTKRSSALRVLPFG							178
Qy	178	SLRLKRNACQQRWYFTFENGACSGPLTIEAIIYLDQSPELNSTINIHRTSSVEGLCEG							237
Db	179	SLRLKRNACQQRWYFTFENGACSGPLTIEAIIYLDQSPENNSTINIHRTSSVEGLCEG							238
Qy	238	IGAGLVDAIVAVWGTCSDYPKGDASTGMSVSRIIITTEELPK							277
Db	239	IGAGLVDAIVAVWGTCSDYPKGDASTGMSVSRIIITTEELPK							278

Search completed: August 4, 2005, 04:45:28  
Job time : 99.1268 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 4, 2005, 04:04:29 ; Search time 106.817 Seconds

(without alignments)  
1002.955 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476  
Sequence: 1 MRPAALGGLTSLRAAGLCRL.....GDASTGMSVSRILIEELPK 277Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1476	100.0	277	5	ABBB80979
2	1476	100.0	277	5	ABG75759
3	1307	88.6	245	5	ABBB80977
4	1307	88.6	245	6	ABG75757
5	1287.5	87.2	244	8	ADOC28703
6	1243.5	84.2	278	3	AAAB08856
7	1243.5	84.2	278	5	ABG96338
8	1243.5	84.2	278	5	ABR47637
9	1240.5	84.0	278	5	ABP68631
10	1239.5	84.0	278	4	AAAM25746
11	1234.5	83.6	278	5	ABP68637
12	1234.5	83.6	278	5	ABG78939
13	1234.5	83.6	278	6	ABJ37762
14	1234.5	83.6	278	7	ADJ93177
15	1225	83.0	243	4	ABBB6667
16	1225	83.0	243	5	ABG96340
17	1225	83.0	243	5	ABJ05554
18	1225	83.0	243	5	ABBB80978
19	1225	83.0	243	6	ABR48554
20	1225	83.0	243	6	ABR48237
21	1225	83.0	243	6	ABG75758
22	1225	83.0	243	6	ABU56607
23	1225	83.0	243	7	ADN39855
24	1225	83.0	243	7	ADN39855
25	1225	83.0	243	7	ADN38732

## ALIGNMENTS

RESULT 1	ADL70254	LBFL301 P
ABBB80979	AA91529	Human sec
ID ABB80979 standard; protein; 277 AA.	ADL71601	Novel hum
XX	ABJ37031	Human bre
AC ABB80979;	AA99442	Human PRO
XX	AA99441	Human PRO
DT 21-OCT-2002 (first entry)	AB66190	Protein O
XX	AA66211	Protein O
DE Long form of rat REMODELIN polypeptide.	AAU9206	Human PRO
XX	AB87586	Human PRO
KM REMODELIN; vulnery; vasotropic; cyostatic; osteopathic; collagen;	ABG95911	Human sec
XX	ABG95911	Human PRO
KW gene therapy; bone; rat.	AB878938	Human bre
XX	ABU58582	Human PRO
OS Rattus sp.	ABU88130	Novel hum
XX	ABU84445	Human sec
PN WO200242487-A2.	ABR66319	Human sec
XX		
PD 30-MAY-2002.		
XX		
Pf 19-OCT-2001; 2001WO-US050940.		
XX		
PR 19-OCT-2000; 2000US-00692081.		
XX		
PA (MAIN-) MAINE MEDICAL CENT RES INST.		
XX		
PI Linder V, Friesel R;		
XX		
DR WPI; 2002-590472/63.		
XX		
DR N-PSDB; ABN6480.		
XX		
PT New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,		
PT negative remodeling, fibrosis, collagen disease, and bone disease in a		
PT mammal.		
XX		
PS Claim 5; Fig 4C; 187dp; English.		
XX		
CC The invention relates to mammalian REMODELIN polypeptides and encoding		
CC polynucleotides. REMODELIN is used to diagnose arterial restenosis,		
CC negative remodeling or fibrosis, bone disease such as osteogenesis		
CC imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis		
CC bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are		
CC used to affect cellular gene expression in a mammal, where the cellular		
CC gene is from transforming growth factor (TGF)-beta1, collagen fibrillar,		
CC osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein		
CC 4. The expression of osteopontin is dependent on Cbfa1. Compositions		
CC comprising antisense REMODELIN sequences are useful for treating diseases		
CC mediated by abnormal expression of a REMODELIN molecule in a human such		
CC as impaired wound healing, fibrosis of an organ, ectopic ossification, or		



XX 21-OCT-2002 (first entry)  
 DT Rat REMODELIN polypeptide.  
 XX REMODELIN; vulnery; vasotrophic; cytostatic; osteopathic; collagen;  
 KM gene therapy; bone; rat.  
 XX Rattus sp.  
 OS WO200242487-A2.  
 PN 30-MAY-2002.  
 PD 19-OCT-2001; 2001WO-US050940.  
 PF 19-OCT-2000; 2000US-00692081.  
 PR (MAIN-) MAINE MEDICAL CENT RES INST.  
 PA Linder V, Friesel R;  
 PI WPI; 2002-590472/63.  
 DR N-PSDB; ABN86480.  
 XX New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
 PT negative remodeling, fibrosis, collagen disease, and bone disease in a  
 PT mammal.  
 XX Claim 5; Fig 4B; 187pp; English.  
 XX The invention relates to mammalian REMODELIN polypeptides and encoding  
 CC polynucleotides. REMODELIN is used to diagnose arterial restenosis,  
 CC negative remodeling or fibrosis, bone disease such as osteogenesis  
 CC imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis  
 CC bullosa (DEB) and Bechlem myopathy in a mammal. The polynucleotides are  
 CC used to affect cellular gene expression in a mammal, where the cellular  
 CC gene is from transforming growth factor (TGF)-beta1, collagen I(alpha1),  
 CC osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein  
 CC 4. The expression of osteopontin is dependent on Cbfa1. Compositions  
 CC comprising antisense REMODELIN sequences are useful for treating diseases  
 CC mediated by abnormal expression of a REMODELIN molecule in a human such  
 CC as impaired wound healing, fibrosis of an organ, ectopic ossification, or  
 CC hypertrophic scar formation. REMODELIN is useful in the development of  
 CC bone during mammalian embryogenesis, proliferation and/or migration,  
 CC cellular signaling, adventitial fibrosis, negative remodeling and arterial  
 CC restenosis, smooth muscle cell proliferation, and arterial remodeling and  
 CC useful in gene therapy. REMODELIN is useful for elucidating the function  
 CC of REMODELIN molecules in a cell, to identify a compound that affects  
 CC REMODELIN expression and/or TGF-beta signaling, as a potential  
 CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
 CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
 CC formation, to promote bone fracture healing, and to increase or decrease  
 CC expression of REMODELIN in mammal. It is also useful for producing  
 CC recombinant cell and transgenic non-human mammals which are useful tools  
 CC for the study of REMODELIN action, for identifying novel diagnostics and  
 CC therapeutics for treatment, and for elucidating the cellular roles of  
 CC REMODELIN. The present sequence represents a rat REMODELIN polypeptide  
 XX  
 SQ Sequence 245 AA;  
 Query Match 88.6%; Score 1307; DB 5; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 1,9e-122;  
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 33 MHPOGRASPOLLLGLFVLLVLLQLQSPASSASNPVKOKALIRORVVDLYNGKLCQG 92  
 DB 1 MHPOGRASPOLLLGLFVLLVLLQLQSPASSASNPVKOKALIRORVVDLYNGKLCQG 60  
 QY 93 PAVGPGRDGPANGIPGTPIGIRGDFKGEKCECLRESFEESWTNYYKQSSSINYGI 152  
 DB 61 PAVGPGRDGPANGIPGTPIGIRGDFKGEKCECLRESFEESWTNYYKQSSSINYGI 120

QY 153 DLGKIAECTPTKMRNSALRVLPSSGLRLKCRNACCCORVYFEPNGACGGPLFEATLYL 212  
 DB 121 DLGKIAECTPTKMRNSALRVLPSSGLRLKCRNACCCORVYFEPNGACGGPLFEATLYL 180  
 QY 213 DQSPPELNGSTINIRHTSVYGLCEGIGAGLVDAVAVWGTCSDYPKDASTGMNSVSR111 272  
 DB 181 DQSPPELNGSTINIRHTSVYGLCEGIGAGLVDAVAVWGTCSDYPKDASTGMNSVSR111 240  
 QY 273 EELPK 277  
 DB 241 EELPK 245  
 RESULT 4  
 ABG75757  
 ID ABG75757 standard; protein; 245 AA.  
 AC ABG75757;  
 XX  
 DT 28-APR-2003 (first entry)  
 XX  
 XX Rat REMODELIN protein.  
 DE  
 XX Rat; adventitia-inducible; REMODELIN; REMODEL;  
 KM transforming growth factor beta; adventitia; vascular remodeling;  
 KM restenosis; vascular injury; antisense therapy; TGF-beta signaling;  
 KM TGF-beta receptor type II; arterial remodeling; bone formation;  
 KM cartilage formation; osteogenesis imperfecta; Bechlem myopathy;  
 KM dystrophic epidermolysis bullosa; negative remodeling; wound healing;  
 KM arterial stenosis; fibrosis; calcification; transplant; heart valve  
 KM heart valve transplant; osteopathic; antiarteriosclerotic; vulnery.  
 XX  
 XX Rattus sp.  
 OS  
 XX US2002161211-A1.  
 PN 31-OCT-2002.  
 PD 19-OCT-2001; 2001US-00045992.  
 PF 19-OCT-2000; 2000US-00692081.  
 PR (LIND/) LINDNER V.  
 PA (FRIE/) FRIESEL R E.  
 PI Linder V, Friesel R;  
 DR WPI; 2003-238238/23.  
 DR N-PSDB; ABX11340.  
 XX New isolated REMODELIN nucleic acid and polypeptide, useful for mediating  
 PT arterial remodeling, formation of bone and cartilage, and the diagnosis  
 PT and treatment of disorders associated with aberrant expression of  
 PT REMODELIN.  
 XX Claim 3; Fig 4B; 81pp; English.  
 PS  
 XX The invention discloses an isolated nucleic acid encoding a mammalian  
 CC adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL  
 CC and/or adventitia induced bone expressed molecule, AIBER). REMODELIN  
 CC expression was induced by transforming growth factor beta (TGF-beta),  
 CC which is important because proliferative events occurring in the  
 CC adventitia contribute to vascular remodeling and restenosis in response  
 CC to vascular injury and TGF-beta has been shown to be a factor involved in  
 CC this. Also disclosed is an antibody raised against REMODELIN and methods  
 CC for treating a disease mediated by abnormal expression of a REMODELIN in  
 CC a human, for identifying a compound that affects or reduces expression of  
 CC REMODELIN in a cell (e.g. antisense therapy), for identifying a compound  
 CC that affects TGF-beta signaling and for increasing or reducing REMODELIN  
 CC expression in a mammal, comprising administering a REMODELIN expression  
 CC increasing or reducing TGF-beta to the mammal, thereby increasing  
 CC REMODELIN expression or inhibiting signalling via the TGF-beta receptor





```
Db      1 MOPAAASRGADADHVPLGLRLQLRAARQPGAMRPOGPAASPORLRG--LILLILLQ 58
QY      58 LSAPSSASENPVKOKALIRQREVVDLYNGMCLQSPAGVPGRGSGANGIPGTGPIGR 117
Db      59 LPAPSSASEIPKQKOKAQLRQREVVDLYNGMCLQSPAGVPGRGSGANGIPGTGPIGR 118
QY      118 DGFKEGKEGCLRESFEESWTPNYKQCSWSLNYGIDLKIAECTPTKRSNSALRVLPFG 177
Db      119 DGFKEGKEGCLRESFEESWTPNYKQCSWSLNYGIDLKIAECTPTKRSNSALRVLPFG 178
QY      178 SLRLKCRNACCQRMWTFENGAECSGPLPIEAILIYDQSPPELNTSTINIHRTSSVEGLCEG 237
Db      179 SLRLKCRNACCQRMWTFENGAECSGPLPIEAILIYDQSPPELNTSTINIHRTSSVEGLCEG 238
QY      238 IGAGLVDAIWMGTCSDPYKGDASTGMNSVSRRIIEELPK 277
Db      239 IGAGLVDAIWMGTCSDPYKGDASTGMNSVSRRIIEELPK 278

RESULT 8
ABR47627
ID      ABR47627 standard; protein; 278 AA.
XX
AC      ABR47627;
XX
DT      12-JUN-2003 (first entry)
XX
DE      Breast cancer associated protein sequence SEQ ID NO:496.
XX
KW      Human; breast cancer; cytostatic; gene therapy.
XX
OS      Homo sapiens.
XX
PN      MO2003004969-A2.
XX
PD      16-JAN-2003.
XX
PF      21-JUN-2002; 2002MO-US019669.
XX
PR      21-JUN-2001; 2001US-0299887P.
PR      27-JUN-2001; 2001US-0301572P.
PR      18-JUL-2001; 2001US-0306501P.
PR      25-SEP-2001; 2001US-0325002P.
PR      05-MAR-2002; 2002US-0362585P.
PR      14-MAY-2002; 2002US-0380391P.
XX
PA      (MILL-) MILLENIUM PHARM INC.
XX
PI      Lillie J, Gannavarapu M, Glatc K, Hoeresh S, Kamatkar S;
PI      Mertens M, Monahan JF, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
PI      Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX
DR      WPI; 2003-210381/20.
XX
DR      N-PSDB; ACC50329.
XX
PT      Breast cancer diagnosis or treatment by comparing the level of expression
PT      of a marker in a patient sample with that in the control non-breast
PT      cancer sample.
XX
PS      Claim 1; SEQ ID NO 496; 128bp; English.
XX
CC      The present invention describes a method for assessing whether a patient
CC      is afflicted with breast cancer. The method comprises comparing a level
CC      of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC      ABR47386 to ABR47632) in a patient sample and the normal level of
CC      expression of the marker in a control non-breast cancer sample, where a
CC      significant increase in the level of expression of the marker in the
CC      patient sample and the normal level is an indication that the patient is
CC      afflicted with breast cancer. The breast cancer associated sequences from
CC      the present invention have cytostatic activities and can be used in gene
CC      therapy. The method is useful for diagnosing and treating breast cancer.
CC      N.B. The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
```

```
CC      at ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 278 AA;
XX
Query Match      84.2%; Score 1243.5; DB 6; Length 278;
Best Local Similarity 86.8%; Pred. No. 5,5e-116;
Matches 243; Conservative
XX
QY      1 MRPAEEL---GQTLSRAGLCRPLCLLCAQSLPHTHMQGPAASQQLGLFLVYLLILLQ 57
Db      1 MOPAAASRGADADHVPLGLRLQLRAARQPGAMRPOGPAASPORLRG--LILLILLQ 58
QY      58 LSAPSSASENPVKOKALIRQREVVDLYNGMCLQSPAGVPGRGSGANGIPGTGPIGR 117
Db      59 LPAPSSASEIPKQKOKAQLRQREVVDLYNGMCLQSPAGVPGRGSGANGIPGTGPIGR 118
QY      118 DGFKEGKEGCLRESFEESWTPNYKQCSWSLNYGIDLKIAECTPTKRSNSALRVLPFG 177
Db      119 DGFKEGKEGCLRESFEESWTPNYKQCSWSLNYGIDLKIAECTPTKRSNSALRVLPFG 178
QY      178 SLRLKCRNACCQRMWTFENGAECSGPLPIEAILIYDQSPPELNTSTINIHRTSSVEGLCEG 237
Db      179 SLRLKCRNACCQRMWTFENGAECSGPLPIEAILIYDQSPPELNTSTINIHRTSSVEGLCEG 238
QY      238 IGAGLVDAIWMGTCSDPYKGDASTGMNSVSRRIIEELPK 277
Db      239 IGAGLVDAIWMGTCSDPYKGDASTGMNSVSRRIIEELPK 278

RESULT 9
ABP66631
ID      ABP66631 standard; protein; 278 AA.
XX
AC      ABP66631;
XX
DT      14-JAN-2003 (first entry)
XX
DE      Human pancreatic cancer expressed protein SEQ ID NO 4551.
XX
KW      Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW      cytostatic; tumour.
XX
OS      Homo sapiens.
XX
PN      MO200260317-A2.
XX
PD      08-AUG-2002.
XX
PF      30-JAN-2002; 2002MO-US002781.
XX
PR      30-JAN-2001; 2001US-0265305P.
PR      31-JAN-2001; 2001US-0265682P.
PR      09-FEB-2001; 2001US-0267568P.
PR      21-MAR-2001; 2001US-0276651P.
PR      28-APR-2001; 2001US-0287112P.
PR      16-MAY-2001; 2001US-0291631P.
PR      12-JUN-2001; 2001US-0305484P.
PR      20-AUG-2001; 2001US-0313992P.
PR      27-NOV-2001; 2001US-0333626P.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Benson DR, Kloe MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
DR      WPI; 2002-627435/67.
XX
DR      N-PSDB; ABV99139.
XX
PT      New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT      diagnosing, preventing and/or treating cancer, particularly pancreatic
PT      cancer.
XX
PS      Claim 2; SEQ ID NO 4551; 300bp + Sequence Listing; English.
XX
XX
```

CC The invention relates to an isolated polynucleotide (I) comprising: (a)  
CC any of a group of over 4000 nucleotide sequences (ABY94628-ABY99145); (b)  
CC complements of (a); (c) sequences consisting of at least 20 contiguous  
CC residues of (a); (d) sequences that hybridize to (a), under moderately  
CC stringent conditions; (e) sequences having at least 75% or 90% identity  
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-  
CC ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer  
CC in a patient and compositions comprising polypeptides, polynucleotides,  
CC antibodies, fusion proteins, T cell populations and antigen presenting  
CC cells expressing the polypeptide are useful in treating pancreatic cancer  
CC and stimulating an immune response. The polynucleotides can be used as  
CC probes or primers for nucleic acid hybridisation, in the design and  
CC preparation of ribozyme molecules for inhibiting expression of the tumour  
CC polypeptides and proteins in the tumour cells, in vaccines and for gene  
CC therapy. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 278 AA;

Query Match 84.0%; Score 1240.5; DB 5; Length 278;  
Best Local Similarity 86.4%; Pred. No. 1.1e-115;  
Matches 242; Conservative 7; Mismatches 26; Indels 5; Gaps 2;

QY 1 MRPAEL--GQTLNRAGLCRPICLLCASQLPHTMHPOGRAASPOLLLGLPLVLLLLIQ 57  
DB 1 MOPAAASERGGADADHVLIGLRLQLRAKROGAMRPOGPASPORLKG--LLLLLLIQ 58  
QY 58 LSAPSSASSENPKYKOKALRQREVDLNYNMCLOGPAGVGRDGPANGIPETPGIPGR 117  
DB 59 LPAPSSASSEIPKOKOKAQLRQREVDLNYNMCLOGPAGVGRDGPANGIPETPGIPGR 118  
QY 118 DGRGKRGKGLRARSFEESWTPNPKQCSWSLNGIDLGKIAECTPTKMSNSALRYLFGS 177  
DB 119 DGRGKRGKGLRARSFEESWTPNPKQCSWSLNGIDLGKIAECTPTKMSNSALRYLFGS 178  
QY 178 SLRLKRNACCQMYFTFNAGCSGPIPIAIIYLDQSPSELNSTINIHRTSSVEGLCEG 237  
DB 179 SLRLKRNACCERNYFTFNAGCSGPIPIAIIYLDQSPSELNSTINIHRTSSVEGLCEG 238  
QY 238 IGAGLVDAIWTGTCSDYPRKGASTGMNSVSRILIEELPK 277  
DB 239 IGAGLVDAIWTGTCSDYPRKGASTGMNSVSRILIEELPK 278

RESULT 10

AAM25746

AAM25746 standard; protein; 278 AA.

AC AAM25746;

DT 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:1261.

XX Human; cancer; HIV infection; human immunodeficiency virus;  
XX antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
XX antibacterial; endocrine; cardiant; central nervous system; vitruide;  
XX anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
XX antilagregant; haemostatic; vulnary; anticancer; osteopathic; eczema;  
XX dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
XX neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
XX immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
XX antianaemytic; rheumatoid arthritis; septic shock; pancreatitis;  
XX cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
XX genetic disease; haematopoietic disorder; platelet disorder; asthma;  
XX thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
XX allergic rhinitis; diabetes; multiple sclerosis; depression;  
XX Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
XX neurological disorder.

XX Homo sapiens.

XX OS

XX

PN W0200153455-A2.

PD 26-JUL-2001.

PF 22-DEC-2000; 2000WO-US035017.

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI, 2001-457603/49.

DR N-PSDB; AAH99687.

PS Claim 20; Page 260; 1217BP; English.

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to

CC AAM25963. The proteins can have activities based on the tissues and cells

CC they are expressed in, such as: antiinflammatory; antirheumatic;

CC antirheumatic; immunosuppressive; antibacterial; endocrine; cardiant;

CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;

CC cardiovascular; antianaemic; antilagregant; haemostatic; vulnary;

CC anticancer; osteopathic; dermatological; antiallergic; antiasthmatic;

CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;

CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides

CC encoding them can be used in gene therapy, antisense therapy and vaccine

CC production. The proteins and polynucleotides are useful for screening for

CC agonists or antagonists of a protein and for the treatment and diagnosis

CC of disorders associated with the activity of a protein e.g. inflammation,

CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,

CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal

CC infections, autoimmunity, genetic diseases, haematopoietic disorders,

CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,

CC osteoporosis, severe combined immunodeficiency, eczema, allergic

CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,

CC Alzheimer's disease, Parkinson's disease, neurodegenerative and

CC neurological disorders

SQ Sequence 278 AA;

Query Match 84.0%; Score 1239.5; DB 4; Length 278;

Best Local Similarity 90.9%; Pred. No. 1.4e-115;

Matches 239; Conservative 5; Mismatches 14; Indels 5; Gaps 2;

QY 15 GLCRPLCLLCSQULPHTMHPOGRAASPOLLLGLPLVLLLLIQLSAPSSASSENPKYKOKA 74  
DB 21 GLLR--IQLRARARPOGAMRPOGPASPORLKG--LLLLLLIQLPAPSSASSEIPKOKOKA 75  
QY 75 LQREVDLNYNMCLOGPAGVGRDGPANGIPETPGIPGRDGRGKGLRARSFE 134  
DB 76 QUREVDLNYNMCLOGPAGVGRDGPANGIPETPGIPGRDGRGKGLRARSFE 135  
QY 135 SWTPNPKQCSWSLNGYDGLGKIAECTPTKMSNSALRYLFGSGSLRLKRNACCQMYFT 194  
DB 136 SWTPNPKQCSWSLNGYDGLGKIAECTPTKMSNSALRYLFGSGSLRLKRNACCQMYFT 195  
QY 195 FNGABCSGGLPIEAIIYLDQSPSELNSTINIHRTSSVEGLCEGIGLVDAIWTGTCSD 254  
DB 196 FNGABCSGGLPIEAIIYLDQSPSELNSTINIHRTSSVEGLCEGIGLVDAIWTGTCSD 255  
QY 255 YPKGDASTGMNSVSRILIEELPK 277  
DB 256 YPKGDASTGMNSVSRILIEELPK 278

RESULT 11

ABP68637

ID	ABP68637	standard; protein; 278 AA.
XX	XX	
XX	AC	ABP68637;
XX	DT	14-JAN-2003 (first entry)
DE	XX	Human pancreatic cancer expressed protein SEQ ID NO 4560.
XX	KW	Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
XX	KW	cytostatic; tumour.
OS	XX	Homo sapiens.
XX	XX	
XX	XX	WO200260317-A2.
XX	XX	
XX	XX	08-AUG-2002.
XX	XX	
XX	XX	30-JAN-2002; 2002W0-US002781.
XX	XX	
XX	XX	30-JAN-2001; 2001US-0265305P.
XX	XX	31-JAN-2001; 2001US-0265682P.
XX	XX	09-FEB-2001; 2001US-0267568B.
XX	XX	21-MAR-2001; 2001US-0279651P.
XX	XX	28-APR-2001; 2001US-0287112P.
XX	XX	16-MAY-2001; 2001US-0291631P.
XX	XX	12-JUL-2001; 2001US-0305484P.
XX	XX	20-AUG-2001; 2001US-0313999P.
XX	XX	27-NOV-2001; 2001US-0333626P.
XX	XX	
XX	XX	(CORI-) CORIXA CORP.
XX	XX	
XX	XX	Benson DR, Kaios MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX	XX	WPI; 2002-627435/67.
XX	XX	DR
XX	XX	N-PSDB; ABV99145.
XX	XX	
PT	PT	New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT	PT	diagnosing, preventing and/or treating cancer, particularly pancreatic
PT	PT	cancer.
PS	PS	Claim 2; SEQ ID NO 4560; 300bp + Sequence Listing; English.
XX	XX	
XX	XX	The invention relates to an isolated polynucleotide (I) comprising: (a)
XX	XX	any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
XX	XX	complements of (a); (c) sequences consisting of at least 20 contiguous
XX	XX	residues of (a); (d) sequences that hybridize to (a), under moderately
XX	XX	stringent conditions; (e) sequences having at least 75% or 90% identity
XX	XX	to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
XX	XX	ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
XX	XX	in a patient and compositions comprising polypeptides, polynucleotides,
XX	XX	antibodies, fusion proteins, T cell populations and antigen presenting
XX	XX	cells expressing the polypeptide are useful in treating pancreatic cancer
XX	XX	and stimulating an immune response. The polynucleotides can be used as
XX	XX	probes or primers for nucleic acid hybridisation, in the design and
XX	XX	preparation of ribozyme molecules for inhibiting expression of the tumour
XX	XX	polypeptides and proteins in the tumour cells, in vaccines and for gene
XX	XX	therapy. Note: The sequence data for this patent did not form part of the
XX	XX	printed specification, but was obtained in electronic format directly
XX	XX	from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	XX	
XX	XX	Sequence 278 AA;
XX	XX	
XX	XX	Query Match 83.6%; Score 1234.5; DB 5; Length 278;
XX	XX	Best Local Similarity 86.4%; Pred. No. 4.4e-115;
XX	XX	Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2
QY	QY	1 MRPAAL---GQTLSDGLCRPLCTLLCASQLPHTNHPOGRAASPOLLLGLFLVLLLLQ 57
DB	DB	1 MQPAASERGGADADHVPILGLRLRLQLRARQPGAMRPGQPAASPRLRG--TLTLLLLQ 58
QY	QY	58 LSAPSSASNPVKYKQKALLRQREVDLLNGMCLQSGAGVPGKRDGSGANGIPETPGIPGR 117
DB	DB	59 LPAPSSASIPGKQKQALRQREVDLLNGMCLQSGAGVPGKRDGSGANGVPTPTPSPG 118

Qy	118	DGRKGEKGLNRSEFEESMTPEYKQCCSSSLNYGIDLGKIAIECTFTKRRNSALRV.FSG	177
Db	119	DGRKGEKGEGLNRSEFEESWTPYKQCCSSSLNYGIDLGKIAIECTFTKRRNSALRV.FSG	178
Qy	178	SLFLKRRNACCORWYTFPGAGCSGPLPIEALIIYDQSPFLNNTINIHRTSVSVEGLCEG	237
Db	179	SLFLKRRNACCORWYTFPGAGCSGPLPIEALIIYDQSPFLNNTINIHRTSVSVEGLCEG	238
Qy	238	IGAGLVDAIIVWGTCSDYPKDGASTGMNSVSRIIIEELPK	277
Db	239	IGAGLVDAIIVWGTCSDYPKDGASTGMNSVSRIIIEELPK	278
RESULT 12			
ABG78939	ID	ABG78939 standard; protein; 278 AA.	
ABG78939	AC	ABG78939;	
XX	DT	15-NOV-2002 (first entry)	
XX	DE	Human breast tumour polypeptide #30.	
XX	DE	Human breast tumour polypeptide #30.	
KW	XX	Human; breast tumour protein; breast cancer; cytostatic; vaccine.	
XX	OS	Homo sapiens.	
XX	PN	US2002085998-A1.	
XX	PD	04-JUL-2002.	
XX	PF	13-APR-2001; 2001US-00834759.	
XX	PR	28-DEC-1998; 98US-00222575.	
XX	PR	02-APR-1999; 99US-00285480.	
XX	PR	23-JUN-1999; 99US-00339338.	
XX	PR	02-SEP-1999; 99US-00389681.	
XX	PR	03-NOV-1999; 99US-00433826.	
XX	PR	17-APR-2000; 2000US-00551621.	
XX	PR	08-JUN-2000; 2000US-00590751.	
XX	PR	22-JUN-2000; 2000US-00604287.	
XX	PR	20-JUL-2000; 2000US-00620405.	
PA		(CORI-) CORIXA CORP.	
PI		Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;	
PI		Henderson RA;	
XX		DR N-PSDB; AB664030, AB664031.	
XX		WPI; 2002-635657/68.	
XX		PT Novel breast cancer polynucleotides and polypeptides encoded by the	
XX		polynucleotides, useful for detecting the presence of breast cancer in a	
XX		patient, and in pharmaceutical compositions, for treating breast cancer.	
XX		PS Claim 2; Page 236-237; 247pp; English.	
XX		CC The invention relates to an isolated breast tumour polynucleotide and the	
XX		polypeptide it encodes. The polynucleotide and polypeptide are useful for	
XX		detecting the presence of breast cancer in a patient, and in	
XX		pharmaceutical compositions for treating breast cancer. The sequences are	
XX		useful for stimulating an immune response in a patient and can therefore	
XX		be used in production of vaccines. The sequences are also useful for	
XX		detecting the presence of a cancer in a patient, by obtaining a	
XX		biological sample from the patient, contacting the biological sample with	
XX		a composition of the invention and detecting the amount of polynucleotide	
XX		that hybridizes to the sample. This sequence represents a human breast	
XX		tumour polypeptide of the invention	
XX		Sequence 278 AA;	
XX		83.6%; Score 1234.5; DB 5; Length 278;	
XX		Query Match	



Best Local Similarity 86.4%; Pred. No. 4.4e-115;  
Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;

```

QY 1 MRAAEI-----GQTSBAGLCRPLCLTCLCSOHPHMHPCGRASPOLLGEFLVLLILO 57
D 1 MGRAPASBREGADADVPLLGLRLQDRAARGPAMRPGRASQRLRG--LILLILLQ 58
QY 58 LSAPSSASENPKYKOKALIRQREVDVLYNMGCLQGPAGYPCGDSPGANGIPGTGPIGR 117
D 59 LPAPSSASELPKKQKQAKOIRQREVDVLYNMGCLQGPAGYPCGDSPGANVLPGTGPIGR 118
QY 118 DGFQKGEAGCELRSPESPSWTPYNYKQCSWSLNYGIDLKVIACCTTKRSNSALRYVFSG 177
D 119 DGRGGEAGCELRSPESWTPYNYKQCSWSLNYGIDLKVIACCTTKRSNSALRYVFSG 178
QY 178 SLRLKCNACCCQRYWTFPFGAECSGPLPIEALITYLDQSPELNSTINIHTRSVYGLCEG 237
D 179 SLRLKCNACCCQRYWTFPFGAECSGPLPIEALITYLDQSPENNSTINIHTRSVYGLCEG 238
QY 238 IGAGLVDAVIAWGTGSCDYPKGDASTGMSVSRILILEELPK 277
D 239 IGAGLVDAVIAWGTGSCDYPKGDASTGMSVSRILILEELPK 278

```

## RESULT 13

ID ABJ37762 standard; protein; 278 AA.

AC	ABJ37762;
XX	
XX	15-MAY-2003 (first entry)
XX	
XX	Human tumour-related protein - SEQ ID No 515.
DE	
XX	Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW	tumour; breast cancer; cancer; immune response stimulation.
OS	Homo sapiens.
XX	
XX	WO200283956-A1.
PN	
PD	24-OCT-2002.
XX	
PF	15-APR-2002; 2002WO-US012378.
XX	
XX	13-APR-2001; 2001US-00834759.
PR	07-DEC-2001; 2001US-00007805.
PR	13-FEB-2002; 2002US-00076622.
XX	
PA	(CORI-) CORIAX CORP.
PI	Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI	Mitcham JL, Xu J, Hallock SL, Hepler WT, Henderson RA, Fanger GR,
PI	Vedvick TS, McNeill PD, Durham M;
XX	
DR	WPI; 2003-103376/09.
XX	
XX	New polypeptide and polynucleotide useful for stimulating and/or
PT	expanding T cells specific for a tumor protein and treating breast
PT	cancer.
XX	
XX	Example 1; Page 322; 375pp; English.
XX	
CC	The invention comprises a method of stimulating and/or expanding T cells
CC	specific for a tumour protein. The invention further comprises human
CC	nucleic acids and proteins that are associated with tumours (e.g. breast
CC	cancer). The method and sequences of the invention are useful for
CC	stimulating and/or expanding T cells specific for a tumour protein,
CC	detecting the presence of cancer, stimulating an immune response in a
CC	patient and treating breast cancer. The present amino acid sequence
CC	represents a human tumour-related protein
XX	
XX	Sequence 278 AA;
XX	

Query Match	83.6%;	Score 1234.5;	DB 6;	Length 278;
Best Local Similarity	86.4%;	Pred. NO. 4.4e-115;		
Matches 242;	Conservative 6;	Mismatches 27;	Indels 5;	Gaps 2;

Oy	1	MRPAAL---	GGTLSPAGCPRPLCLLCLASOLPHFMHQGAAPOLLGLFVILLQ	57
Db	1	MGPAASASERGDADHVPLPLGLPLRLQDLAAHQPGAMRQGPAAAPQLRAG--	LLLLLLQ	58
Oy	58	LSAPSSASENPKYKQKALIRQREVDYLNGMCLQGPAGVPGRDSPGANGIPGTPGIPGR	117	
Db	59	LPPASSASBIPKGGKQKQALQRLREVDYLNGMCLQGPAPVPRDSDPGANVPGTPGIPGR	118	
Oy	118	DGKSGKGECLRESFEESWTPTYKXOCSSSLNYGLDGLKIAECTFTKMRSSALRVLPSG	177	
Db	119	DGKSGKGECLRESFEESWTPTYKXOCSSSLNYGLDGLKIAECTFTKMRSSALRVLPSG	178	
Oy	178	SLFLKCRNAACCRWYFTFGAECSPDLIEAIIYLDQSPBLNSTINIHRTSSVEGLCEG	237	
Db	179	SLFLKCRNAACCRWYFTFGAECSPDLIEAIIYLDQSPBLNSTINIHRTSSVEGLCEG	238	
Oy	238	IGAGLVDAIWIWGTGSDYPKGDASGNNVSARIIIBELPK	277	
Db	239	IGAGLVDAIWIWGTGSDYPKGDASGNNVSARIIIBELPK	278	

## RESULT 14

ID ADL93177 standard; protein; 278 AA.

AC	ADD93177,
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Human breast cancer-associated polypeptide #27.
XX	
KW	gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX	
OS	Homo sapiens.
XX	
PN	US2003166022-A1.
XX	
PD	04-SEP-2003.
XX	
PF	15-APR-2002; 2002US-00124805.
XX	
PR	28-DEC-1998; 98US-00222575.
XX	
PR	02-APR-1999; 99US-00285480.
XX	
PR	23-JUN-1999; 99US-00339338.
XX	
PR	02-SEP-1999; 99US-00389681.
XX	
PR	03-NOV-1999; 99US-00433826.
XX	
PR	17-APR-2000; 2000US-00551621.
XX	
PR	08-JUN-2000; 2000US-00590751.
XX	
PR	22-JUN-2000; 2000US-00604287.
XX	
PR	10-JUL-2000; 2000US-00620405.
XX	
PR	13-APR-2001; 2001US-00834759.
XX	
PR	07-DEC-2001; 2001US-00007805.
XX	
PR	13-FEB-2002; 2002US-00076622.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Houghton RL, Sleath PR, Persing DH;
XX	
DR	WPI, 2003-874918/81.
XX	
DR	N-PSDB; ADD93174.
XX	
PT	An isolated oncogenic polypeptide useful for preventing, diagnosing and
XX	treating breast cancer.
PS	Example 1; SEQ ID NO 515; 294pp; English.
CC	The invention relates to an isolated breast cancer-associated
CC	polypeptide. The polypeptide may be used for the diagnosis and treatment



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 4, 2005, 04:07:24 ; Search time 101.024 Seconds

(without alignments)  
1404.088 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476

Sequence: 1 MRPAALGQTLRSAGLCRPL.....GDASTGMSVSRILIELPK 277

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot\_03:\*  
2: uniprot\_sprot:\*  
3: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1307	88.6	245	2	08CG08
2	1292	87.5	245	2	08DND6
3	1225	83.0	243	2	096CG8
4	1216	82.4	243	2	06UW91
5	1063	72.0	232	2	081X63
6	932	63.1	231	2	06AXL0
7	133	9.0	287	2	07Z513
8	133	9.0	287	2	08CFR0
9	132.5	8.9	555	2	08K036
10	131.5	8.9	312	2	08CHX9
11	130.5	8.8	289	2	08BVD7
12	130.5	8.8	551	2	06ZM13
13	129.5	8.8	1669	2	09QZS0
14	127.5	8.6	289	2	018799
15	126	8.5	305	2	09UDP6
16	125.5	8.5	291	2	09NAR3
17	125.5	8.5	298	2	018286
18	125.5	8.5	1752	2	007265
19	125	8.5	717	2	09N052
20	125	8.5	1163	2	08N6U4
21	125	8.5	1466	1	CA13 HUMAN
22	124.5	8.4	240	2	06LA35
23	124.5	8.4	326	2	0677V9
24	124.5	8.4	358	2	06MEY7
25	124.5	8.4	1336	2	06RZ41
26	124.5	8.4	1516	1	CA1H HUMAN
27	124.5	8.4	1516	2	06RZ39
28	124.5	8.4	1751	2	06RZ40
29	123.5	8.4	540	2	001904
30	123.5	8.4	957	2	086Y22
31	123.5	8.4	957	2	096P44

32	123.5	8.4	957	2	09H0V3	09h0v3 homo sapien
33	123	8.3	1472	2	090ZAO	090zao gallus gall
34	122.5	8.3	282	2	08WFP6	08wfp6 suberites d
35	122.5	8.3	310	2	020282	020282 caenorhabdi
36	122.5	8.3	546	2	06P7U1	06p7u1 mus musculu
37	122.5	8.3	920	2	078EC6	078ec6 mus sp. typ
38	122.5	8.3	1549	2	060444	060444 cricetus
39	122.5	8.3	2944	2	063870	063870 mus musculu
40	122	8.3	178	2	07M2W9	07m2w9 bos taurus
41	122	8.3	549	2	08EMF8	08emf8 mus musculu
42	122	8.3	747	1	CA12 BOVIN	P02459 bos taurus
43	122	8.3	1284	2	06P159	06p159 homo sapien
44	121.5	8.2	289	1	COT7 HUMAN	09bx12 homo sapien
45	121	8.2	755	2	000261	000261 homo sapien

## ALIGNMENTS

RESULT 1	ID	Q8CG08	PRELIMINARY;	PRT;	245 AA.
AC	Q8CG08				
DT	01-MAR-2003 (Tremblrel. 23, Created)				
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)				
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)				
DE	Collagen triple helix repeat-containing protein 1.				
GN	Name=Cthrc1;				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Sprague-Dawley; TISSUE=8 day balloon-injured carotid artery;				
RA	Lehnert W., Moore D.P., Harmon K.U., Mancini M.L., Lindner V.;				
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AY136824; AAN15748.1; -				
DR	GO; GO:0005737; Cytoplasm; IEA.				
DR	GO; GO:0006817; P:phosphate transport; IEA.				
DR	InterPro; IPR008161; C1g helix.				
DR	InterPro; IPR008160; Collagen.				
DR	Pfam; PF01391; Collagen; 1.				
DR	ProDom; PD000007; C1g_helix; 1.				
KW	Collagen.				
SQ	SEQUENCE 245 AA; 26424 MW; 2296FD6DCDBA21F2 CRC64;				
Query Match	88.6%; Score 1307; DB 2; Length 245;				
Best Local Similarity	100.0%; Pred. No. 1.8e-106;				
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	33 MHQGRASPOLLGLFLVLLLLQLSAPSSASSENKVKOKLIRQEVVDLYNMCLOG 92				
DB	1 MHQGRASPOLLGLFLVLLLLQLSAPSSASSENKVKOKLIRQEVVDLYNMCLOG 60				
QY	93 PAVGPRDSSPRANGIPGTPGIPGRGFGKGEKCELRSEFESWTNYKQCSNLSNYGI 152				
DB	61 PAVGPRDSSPRANGIPGTPGIPGRGFGKGEKCELRSEFESWTNYKQCSNLSNYGI 120				
QY	153 DLGKIACEFTKRSNSALRVLFSGSLRLKCRNACQRYFTFNGAECGSLPIEAIYLY 212				
DB	121 DLGKIACEFTKRSNSALRVLFSGSLRLKCRNACQRYFTFNGAECGSLPIEAIYLY 180				
QY	213 DQSPPELNTINIHRTSVYGLCEGIGAGLVDAIVWGTCSDYPKGDASTGMSVSRIII 272				
DB	181 DQSPPELNTINIHRTSVYGLCEGIGAGLVDAIVWGTCSDYPKGDASTGMSVSRIII 240				
QY	273 EELPK 277				
DB	241 EELPK 245				
RESULT 2					

Q9D1D6  
ID Q9D1D6 PRELIMINARY; PRT; 245 AA.  
AC Q9D1D6;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:1110014B07 product:hypothetical Collagen triple helix repeat containing protein, full insert sequence.  
DE repeat containing protein, full insert sequence.  
GN Name=Ctrnc1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA The FANTOM Consortium.  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Komno H., Akiyama J., Nishi K., Kiteunai T., Taehiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa K., Izawa M., Ohara E., Wachihi K., Yonekura Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanaoka T., Hara A., Hayatsu N., Hiramoto K., Hiraoaka T., Hori F., Imoto K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Komno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Toyota T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK003674; BAB22930.1; -.  
DR WGI; WGI:1915638; Ctrnc1.  
DR GO; GO:0005737; Cytoplasm; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR008161; C1g helix.  
DR InterPro; IPR008160; Collagen.  
DR Pfam; PF01391; Collagen; 1.  
DR ProDom; PD000007; C1g\_helix; 1.  
KW Collagen; Hypothetical protein.  
SQ SEQUENCE 245 AA; 26460 MW; 14951887D8181A0E CRC64;  
Query Match 87.5%; Score 1292; DB 2; Length 245;  
Best Local Similarity 98.8%; Pred. No.3,8e-105;  
Matches 242; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 33 MHPOGRAAPOLLGLFLVLLILQLSAPSSASNPYKQKALIRQREVVDLYNMCLOG 92  
DB 1 MHPOGRAAPOLLGLFLVLLILQLSAPISASEPKYQOALLIRQREVVDLYNMCLOG 60  
QY 93 PAVGVRGRGSPGANGIPGTPTGIPGRDGPKEGEGELRESFESWTPNPKQCSWSLNGI 152  
DB 61 PAVGVRGRGSPGANGIPGTPTGIPGRDGPKEGEGELRESFESWTPNPKQCSWSLNGI 120  
QY 153 DLGKIAECTFTMRNSALRYLFSGSLRLKCRNACCOWYFTPNGAECSPPIPIATYL 212  
DB 121 DLGKIAECTFTMRNSALRYLFSGSLRLKCRNACCOWYFTPNGAECSPPIPIATYL 180  
QY 213 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAVIAWGTCSDPYKGDASTGMNSVRIIT 272  
DB 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAVIAWGTCSDPYKGDASTGMNSVRIIT 240  
QY 273 EELPK 277  
DB 241 EELPK 245  
RESULT 3  
Q96CG8 PRELIMINARY; PRT; 243 AA.  
ID Q96CG8  
AC Q96CG8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Collagen triple helix repeat containing 1 (collagen triple helix repeat-containing protein 1).  
GN Name=CTRNC1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K., Hopkins M., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McSwan P.J., McKernan K.O., Malek J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Whiting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting R.M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywicki M.I., Skalska U., Smailus D.E., Scherch N., Schein J.E., Jones S.J., Meira M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human

Query	Best local similarity	Score 1225;	DB 2;	Length 243;
Matches 232;	Conservative 3;	Mismatches 8;	Indels 2;	Gaps 1
33	MHPQRAASPOLLGLFLVLLTLTLQASASSSENPKVOKALIRQREVVYDLYNGMCLOG	94.7%;	Pred. No. 2,8e-99;	
1	MRPGPAPSPQRLRG-LTLTLTLQLPAPSSASSEIKGQKQALRQREVVYDLYNGMCLOG			58
93	PAGVGRDGPSPGANGIPGTPIGPGRDGFKEGCECLRESPEBSWNTNYQCSWSLNYGI			1522
59	PAGVGRDGPSPGANGIPGTPIGPGRDGFKEGCECLRESPEBSWNTNYQCSWSLNYGI			118
153	DLGKIAECTFTMRNSALRVLPSSGLRLKCRNACCCORNYFTFNAGCGPLIEAIIYL			212
119	DLGKIAECTFTMRNSALRVLPSSGLRLKCRNACCCORNYFTFNAGCGPLIEAIIYL			178
213	DOGSSELNSTIHHSTSVSEVGECEGIGAGLVDAIVWGCSDPKDASTGNMVSARIIT			272
179	DOGSSELNSTIHHSTSVSEVGECEGIGAGLVDAIVWGCSDPKDASTGNMVSARIIT			238
273	EELPK 277			
239	EELPK 243			

RA	"The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment."
RT	Genome Res. 13:2265-2270(2003).
RL	EMBL; AY58914; AAC08273.1; -.
DR	GO; GO:0005737; Cytoplasm; IEA.
DR	GO; GO:000617; P-phosphate transport; IEA.
DR	InterPro; IPR008161; Ctg_helix.
DR	InterPro; IPR008160; Collagen.
DR	Prodomb; PD000007; Ctg_helix; 1.
KW	Collagen.
SQ	SEQUENCE 243 AA; 2626 MW; BCB49AF4DBC303BC CRC64;
OY	Query Match 82.4%; Score 1216; DB 2; Length 243; Best Local Similarity 94.3%; Pred. No. 1.7e-98; Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1
Dd	33 MHFGGAASPOLLLGLFVLLILLQLTSAASSASENPKVKALIRREVVLYLNGMCLQG 1 MRGGRABAPORLRG-LILLULLDLPAPPSSASEIPKQKQAQILRQBVDVLNGMCLQG 58
Oy	93 PACVPGRDSDPGANGIPGTPTGIPGRDGFKGEGCELCRESFEBSWTPTYKOCSSSLNYGI 152
Dd	59 PACVPGRDSDSGANVIPTGPPIGRDGFKGEGCELCRESFEBSWTPTYKOCSSSLNYGI 118
Oy	153 DLCKIAECTFTMRNSALRVFPSSLKLKRNACCQRWYTFMGACSCGPLPEAIITYL 212
Dd	119 DLCKIAECTFTMRNSALRVFPSSLKLKRNACCQRWYTFMGACSCGPLPEAIITYL 178
Oy	213 DGSPEPLNTINIHRTSSVEGLCEGIGAGLDVAIHWGTCSDYPKDGASTGMNSVSRIII 272
Dd	179 DGSPPEMLNTINHRTSSVEGLCEGIGAGLDVAIHWGTCSDYPKDGASTGMNSVSRIII 238
Oy	273 EEELPK 277       239 EEELPK 243
Dd	
RESULT 5	
Q8IX63	PRELIMINARY; PRT; 232 AA. O8IX63
AC	O8IX63;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Laet annotation update)
OS	Homo sapiens (Human).
DE	NTRYC1.
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. OX NCBI_TaxID=9606;
KW	[1]
RN	SEQUENCE FROM N.A.
RP	Sanuki N., Fujiki K., Kanai A., Tanaka Y., Iwata T.; RA Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases. DR EMBL; AF395488; AAC017919.1; -.
DR	GO; GO:0005737; Cytoplasm; IEA.
DR	GO; GO:000617; P-phosphate transport; IEA.
DR	InterPro; IPR008161; Ctg_helix.
DR	InterPro; IPR008160; Collagen.
DR	Pfam; PF01391; Collagen; 1.
DR	Prodomb; PD000007; Ctg_helix; 1.
KW	Collagen.
SQ	SEQUENCE 232 AA; 25163 MW; EBP4BC30304837ED CRC64;
OY	Query Match 72.0%; Score 1063; DB 2; Length 232; Best Local Similarity 87.2%; Pred. No. 4.4e-85; Matches 197; Conservative 7; Mismatches 12; Indels 10; Gaps 1
Dd	61 PSSASNPVKOKALIRRE-----VDLYNMCCAGPVGVRGDSGGANGIRG 110 ::: : 3 PPSRSITVLRKRYSRKLENNGPSAFQGLICGYRNKCCAGPVGVRGDSGGANGIRG 62

Qy 111 TPGIGRGRDGFKEKGEKCHLRESFEESWTPTYKQCSWSLWYGDLDGIAECTPTKMSNSA 170  
 Db 63 TPGIGRGRDGFKEKGEKCHLRESFEESWTPTYKQCSWSLWYGDLDGIAECTPTKMSNSA 122  
 Qy 171 LRVFSSGLRLCKRNACCQRWYFTFNGACSCGLPLEAIIYYLDQSGPELWSTNINIRTS 230  
 Db 123 LRVLFSSGLRLCKRNACCQRWYFTFNGACSCGLPLEAIIYYLDQSGPEMNSTINIRTS 182  
 Qy 231 VEGLEGGTGAAGLVDAIWTGTCSDYKGGASTCWNVSRIIIBELP 276  
 Db 183 VEGLEGGTGAAGLVDAIWTGTCSDYKGGASTCWNVSRIIIBELP 228

Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
52	63.1%	Score 932;	DB 2;	Length 231;		
8	75.4%	Pred. No. 1,4e-73;				
172;	21;	Mismatches	29;	Indels	6;	Gaps
110	63.1%	Score 932;	DB 2;	Length 231;		
8	75.4%	Pred. No. 1,4e-73;				
172;	21;	Mismatches	29;	Indels	6;	Gaps

[illegible]

<b>RESULT 7</b>					
ID	Q7Z5L3	PRELIMINARY;	PRT,	287 AA.	
AC	Q7Z5L3				
DT	01-OCT-2003 (TrEMBLrel. 25, Created)				
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
DE	C1q-domain containing protein.				
OS	Homo sapiens (human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Ding P., Han W., Wang Y., Rui M., Song Q., Zhang Y., Ma D.;				
RL	Submitted (JUN-2002) to the EMBL/genbank/DBJ databases.				
DR	EMBL; AF525315; AAP80787.1; -.				
DR	HSSP; Q60994; IC28.				
DR	GO; GO:0005737; C:cyttoplasm; IEA.				
DR	GO; GO:0006817; P:phosphate transport; IEA.				
DR	InterPro; IPR001073; Clq.				
DR	InterPro; IPR008160; Collagen.				
DR	InterPro; IPR008983; TNF_like.				
DR	Pfam; PF00386; Clq; 1.				
DR	Pfam; PF01391; Collagen; 1.				
DR	SMART; SM00110; Clq; 1.				
DR	PROSITE; PS01113; Clq; 1.				
KW	Collagen.				
SQ	SEQUENCE 287 AA; 29591 MW; EGC21A8D5F8650A CRC64;				
Query Match	9.0%;	Score 133;	DB 2;	Length 287;	
Best Local Similarity	27.3%;	Pred. No. 0.0022;			
Matches	44;	Conservative 18;	Mismatches 67;	Indels 32;	Gaps 2
Oy	44 ILGLFLVLLILLQLS-----APSSASENPVKYKAKI 76				
Dd	1 MALGLIAIVPLLIQAIPRGAAHYEMMGTCRMIEWPYTAPGCGEPFGAKANPGSTAALE 60				
Oy	77 ROREVVDLVNGMCLQGPAVGPRDPSPGANGIPGTPIGIRGRDGKGEGKECTLRSPESW 136				
Dd	61 VMGDLANPPPPRIQPKKDPPRPPKPPGRPPGPDPGGPEKGDGRGL----- 116				
Oy	137 TPNYKCCSWSSLNYGIDLGIACEFTTKRSNSALRVLFSG 177				
Dd	117 -PGLQTAGTAGSAGVGWGAGVGSDGEVTYSALSATPFSG 156				
<b>RESULT 8</b>					
Q8CFRO	PRELIMINARY;	PRT,	287 AA.		
ID	Q8CFRO				
AC	Q8CFRO				
DT	01-MAR-2003 (TrEMBLrel. 23, Created)				
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
DE	Similar to C1q-like.				
GN	Name=BC040774;				
OS	Mus musculus (mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				

RP SEQUENCE FROM N.A.  
RC TISSUE=Eye;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywiński M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Eye;  
RA Strausberg R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC040774; AAH40774.1; -  
DR HSSP; Q60994; IC28.  
DR MGD; MGI:3032521; BC040774.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR CO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR001073; Clq.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR008983; TNF\_like.  
DR Pfam; PF00386; Clq; 1.  
DR DR Pfam; PF01391; Collagen; 1.  
DR PRINTS; PRO0007; COMPLEMENTC1Q.  
DR SMART; SM00110; C1Q; 1.  
DR PROSITE; PS01113; C1Q; 1.  
DR Collagen.  
KW COLLAGEN.  
SQ SEQUENCE 287 AA; 29292 MW; 8FP69EC1C7420415 CRC64;  
  
Query Match 9.0%; Score 133; DB 2; Length 287;  
Best Local Similarity 27.5%; Pred. No. 0.0022;  
Matches 46; Conservative 16; Mismatches 61; Indels 44; Gaps 3;  
  
QY 44 LILGLFVLILLILQLSPSSASSENPKVKALIFOREVVDLYN----- 86  
DB 1 MALGLLIVFLILQLQAAAPGAA-----HYMLGTCTWICDPYVAVAGCPAGAKAPPPGP 54  
QY 87 -----GMCLOGPAGVGRDGPANGIPGTGIPGRDGFKEGEGCLRE 130  
DB 55 STAALEWMDLSANPPPPFQGPGRKGRGKPPGPPEPGRPPGRGPGSGR 114  
QY 131 SEFSEWTPNKKQCSMSLNYGIDLGKIAECTFTKMSNMLRLVBSG 177  
DB 115 GL-----PGLQLTTSAGGVVSGGTGGGGTGEVTSALSAFSG 156  
  
RESULT 9  
O8K036 PRELIMINARY; PRT; 565 AA.  
AC O8K036;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Col13a1 protein.  
GN Name=Col13a1; (Mouse).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=mx FVB/N; TISSUE=Mammary tumor;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywiński M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=mx FVB/N; TISSUE=Mammary tumor;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC034164; AAH34164.1; -  
DR MGD; MGI:1272201; Col13a1.  
DR GO; GO:0005911; C:intercellular junction; IDA.  
DR InterPro; IPR008161; Clq\_helix.  
DR InterPro; IPR008160; Collagen.  
DR Pfam; PF01391; Collagen; 5.  
DR Prodom; PD000007; Clq\_helix; 1.  
KW COLLAGEN.  
SQ SEQUENCE 565 AA; 56726 MW; DBD3FP99D670195F CRC64;  
  
Query Match 9.0%; Score 132.5; DB 2; Length 565;  
Best Local Similarity 32.7%; Pred. No. 0.0052;  
Matches 36; Conservative 10; Mismatches 35; Indels 29; Gaps 5;  
  
QY 59 SAPSSASSENPKVKKALIFOREVVDLYN-----MCLQGPAGVGRDGPANG 107  
DB 402 SGPFGAKGPP-----GKGEHVD-YNGSINEALQEIIRTLALMGPPGIPGOTGPPGPG 452  
QY 108 IP-----CTPGIPGRDGFKEGEGCLRESEFSEWTPNKKQCSMSLNYG 151  
DB 453 TPQQRGEIQLPGPPGHDGKPPRG--KPGDGAGWSSRTPRKGMASRNEG 499  
  
RESULT 10  
O8CHX9 PRELIMINARY; PRT; 312 AA.  
AC O8CHX9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Clgdnf7 protein (Fragment).  
GN Name=Clgdnf7; (Mouse).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,





Oy		22	L I C A S Q P R M H N G C G R A S P O L U G F V L L L L Q L S A P S A S E N P K Y K X A L I R G R E V	81
Dd		10	L A T C A S G P R P R N O A G E S P R Y ----- C I P G -----	38
Oy		82	V D L N G M C L O G P A G V R G R G S P G A N G I G T P G I P R D P K E G E	126
Dd		39	----- L P R P R P G A N G S G R G R G L P R D R D R K E G E	75
 RESULT 12 06ZMT3 PRELMINARY; PRT; 551 AA.				
ID	06ZMT3			
AC	06ZMT3:			
DT	05-JUL-2004 (TREMBLrel. 27, Created)			
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)			
DE	Hypoetical protein FLJ23917.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. [1]			
RN	NBI_TaxID=9606; [1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Colon;			
RA	Kawabara A., Hikiji T., Kobatate N., Inagaki H., Ikema Y., Okamoto S., Okltani R., Oca T., Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.; Submitted (APR-2004) to the EMBL/Genbank/DDBJ databases.			
RL	EMBL; AK12756; BAD18742.1; -			
DR	GO; GO:0005737; Cytoplasm; IEA.			
DR	GO; GO:0006817; Phosphate transport; IEA.			
DR	InterPro; IPR008161; Cig_helix.			
DR	InterPro; IPR008160; Collagen.			
DR	InterPro; IPR003112; Olfac_like.			
Pfam; PF01391; Collagen_2.				
pfam; pf02191; OLF_1.				
Dr	Pfdom; Pf000007; Cig_helix; 1.			
DR	SMART; SMART; OLF_1.			
KW	Collagen.			
SQ	SEQUENCE 551 AA; 58957 MW; CE14A36120DEC18 CRC64;			
Query Match	8.8%; Score 130.5; DB 2; Length 551;			
Best local Similarity	30.7%; Pred. No. 0.0076;			
Matches 35; Conservative 18; Mismatches 20; Indels 41; Gaps 5;				
Oy		35	P G R A S P O L U G F V L L L L Q L S A P S A S E N ----- P K Y K X ----- AL 75	
Dd		76	P R G A S A P Q ----- D P A S A N K K S H S G E P A F H R A S H D M M N T Y S M  ::	119
Oy		76	I R O R E V D L Y N --- G M C L O G P A G V R G R G S P G A N G I G T P G I P R D P K E G E	126
Dd		120	V P R M V D L N G I S T K I C L T G R S G R P ---- G R G A G L P h N G L D G P R G P K E	170
 RESULT 13 09QZSO PRELMINARY; PRT; 1669 AA.				
ID	09QZSO			
AC	09QZSO:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Alpha 3 collagen IV.			
GN	Name=COL4A3;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NBI_TaxID=10090; [1] [1]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Kidney.			
RC	MEDLINE=200005534; PubMed=10534397; DOI=10.1006/geno.1999.5943;			
EX	Lu W., Phillips C.L., Killen P.D., Hlaing T., Harrison W.R., Elser F.F.B., Miner J.H., Overbeek P.A., Meisler M.H.;			

RT	"insertional mutation of the collagen genes Col4a3 and Col4a4 in a mouse model of Alport syndrome."
RL	Genomics 61:113-124(1999).
DR	EMBL; AF169387; AAD50449.1; -.
DR	PIR; I48302; I48302.
DR	HSSP; P02462; ILIL.
DR	MGI; MGI:104688; Col4a3.
DR	GO; GO:0005587; C:collagen type IV; ISS.
DR	GO; GO:0005615; C:extracellular space; TAS.
DR	GO; GO:0005178; F:integrin binding; ISS.
DR	GO; GO:0008191; F:metalloendopeptidase inhibitor activity; ISS.
DR	GO; GO:0006919; P:caspase activation; ISS.
DR	GO; GO:0008283; P:cell proliferation; ISS.
DR	GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; ISS.
DR	GO; GO:0016525; P:negative regulation of angiogenesis; ISS.
DR	InterPro; IPRO08161; Clg helix.
DR	InterPro; IPRO08160; Collagen.
DR	InterPro; IPRO04829; Csauface antigen.
DR	InterPro; IPRO01442; Procollagn_C.
DR	InterPro; IPRO00504; RNA_rec_mot.
DR	Pfam; PF01413; C4; 2.
DR	Pfam; PF01391; Collag; 21.
DR	ProDom; PD000007; clg_helix; 9.
DR	ProDom; PD153432; Csauface_antigen; 2.
DR	ProDom; PD003923; Procollagna_C; 2.
DR	SMART; SMO0111; C4; 2.
DR	PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW	Collagen.
SQ	SEQUENCE 1669 AA; 161769 MW; 30976E59739A47B2 CRC64;
Query Match	8.8%; Score 129.5; DB 2; Length 1669;
Best Local Similarity	37.0%; Pred. No. 0.033;
Matches 37; Conservative 9; Mismatches 29; Indels 25; Gaps 4;	
CY	33 MHPCGRASQLLGLFLVLLILLQLSAPSSASENPKKALIRREVVDLYNGKL-- 90     : 1 MHSK---TAPRFL-VELLTLTLLLAASPASKGCVCCK-----GQCLCA 42
DB	91 -----GGPAGVPGRDSPGANGIPGTGPICRGDGFKEKG 125 : 43 GTKEKEGEKVPSGPFPGQGKFPFPGPLPGPOGPKSPG 82
RESULT 14	
ID Q18799	PRELIMINARY; PRT; 289 AA.
AC Q18799;	
DT 01-NOV-1996 (TrEMBLrel. 01, Created)	
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)	
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE Hypothetical protein C5J3B4.5.	
GN ORFNames=C5J3B4.5;	
OS Caenorhabditis elegans.	
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;	
OC Rhabdilitidae; Peioderinae; Caenorhabditis.	
OX NCBI_TaxId=6239;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=Bristol N2;	
RX MEDLINE=99069613; PubMed=9851916;	
RA none;	
RT "Genome sequence of the nematode C.elegans: A platform for investigating biology";	
RL Science 282:2012-2018(1998).	
RU [2]	
RP SEQUENCE FROM N.A.	
RC STRAIN=Bristol N2;	
RA Berk's M.;	
RL Submitted (JBC-1995) to the EMBL/GenBank/DDBJ databases.	
DR EMBL; Z68215; CA92453.1; --.	
DR PIR; T20177; T20177.	
DR Wormbase; WBGene0000693; C5J3B4.5.	

DR WormRep: CS3B4.5; CE03091.  
DR GO; GO:0005737; Cytoplasm; IEA.  
DR GO; GO:0042302; F:structural constituent of cuticle; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR002486; Col\_cuticle\_N.  
DR Pfam; PF01391; Collagen; 2.  
DR Pfam; PF01484; Col\_cuticle\_N; 1.  
DR Collagen: Hypothetical protein.  
SQ SEQUENCE 289 AA; 28933 MW; 2999A3FF9CC0B044 CRC64;

Query Match 8.6%; Score 127.5; DB 2; Length 289;  
Best Local Similarity 48.1%; Pred. No. 0.0068;  
Matches 25; Conservative 4; Mismatches 8; Indels 15; Gaps 1;

QY 91 QGPAGVPGRDG-----PGANGIPGTPIPGRDGFKEKGEK 127  
DB 220 RGPAGQPKDGAQGKPGKANGEPGQPRDQPRGRPGQGRDGHPEKGV 271

## RESULT 15

ID Q9UDF6 PRELIMINARY; PRT; 305 AA.  
AC Q9UDF6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Type XIV collagen (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCB1\_Taxid=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95128518; PubMed=7827751; DOI=10.1016/0945-053X(94)90194-5;  
RA Brown J.C., Golbik R., Mann K., Timpl R.;  
RL Matrix Biol. 14:287-295(1994).  
DR PIR; S37749; S37749.  
DR PIR; S46557; S46557.  
DR GO; GO:0005737; Cytoplasm; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR008161; Clg\_helix.  
DR InterPro; IPR008160; Collagen.  
DR Pfam; PF01391; Collagen; 4.  
DR ProDom; PD000007; Clg\_helix; 2.  
KW Collagen.  
SQ SEQUENCE 305 AA; 28934 MW; 2F8CF9423B11FDE2 CRC64;

Query Match 8.5%; Score 126; DB 2; Length 305;  
Best Local Similarity 31.9%; Pred. No. 0.0098;  
Matches 37; Conservative 11; Mismatches 34; Indels 34; Gaps 5;

QY 32 TMAPQRAAPQULLFLVLLLLQLSAPSSASENPKVKQKALIR--QREVVD----- 83  
DB 127 SMGPQAGALGPV-----PGAKGERGERDLOQAMVXVAVQVXEQLQSH 173

QY 84 --LYNQM-----CLQGPAGVPGRDGSPGA--NGIPGTPIPGRDGFKEKGE 126  
DB 174 MARIHQIPSHSSIRTVQGPGRPGSPAGPQGPETPGFPGNAGVPTTPE 229

Search completed: August 4, 2005, 04:21:24  
Job time : 102.024 secs



```
QY      63 SASENPKVKOKALIRQREVDLYNGWCLQGPAGVPGRDSPGANGIPGTGIGIRGDFGK 122
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB      82 SSETTRRLRGRDAACQ-----GCCCL---PGLPDPDPGRNGAPRGAPGARPGFPG 130
QY      123 EKGECLRESFESWTP 138
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB      131 RPAVCEETPEPTCTP 146

RESULT 3
A45407
collagen alpha 3(IV) chain - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 22-Sep-1993 #sequence revision 18-Nov-1994 #text change 09-Jul-2004
C:Accession: A45407; A43903; A23940
R:Exposito, J.Y.; D'Alessio, M.; Di Liberto, M.; Ramirez, F.
J. Biol. Chem. 268, 5249-5254, 1993
A:Title: Complete primary structure of a sea urchin type IV collagen alpha chain and
A:Reference number: A45407; MUID:93186842; PMID:8444899
A:Accession: A45407
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1752 <Exp>
A:Cross-references: UNIPROT:Q26312
A:Note: sequence extracted from NCBI backbone (NCBIP:126841)
R:Wessel, G.M.; Etlin, M.; Benson, S.
Dev. Biol. 148, 261-272, 1991
A:Title: Primary mesenchyme cells of the sea urchin embryo require an autonomously produ
A:Reference number: A43903; MUID:92038439; PMID:1936564
A:Accession: A43903
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 'P', 633-1537, 'G' <MES>
A:Cross-references: GB:S64572; NID:q238616; PIDN:AA820270.1; PID:q238617
A:Note: sequence extracted from NCBI backbone (NCBIN:64572, NCBIP:64573)
R:Venkatesan, M.; De Pablo, F.; Vogell, G.; Simpson, R.T.
Proc. Natl. Acad. Sci. U.S.A. 83, 3351-3355, 1986
A:Title: Structure and developmentally regulated expression of a Strongylocentrotus purp
A:Reference number: A23940; MUID:86205894; PMID:3458186
A:Accession: A23940
A:Molecule type: DNA
A:Residues: 742-812 <VEN>
A:Cross-references: EMBL:M13206
C:Superfamily: collagen alpha 1(IV) chain
C:Keyword: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F:162-1532/Region: interrupted helical
F:1524-1752/Domain: carboxyl-terminal nonhelical, NCI <NC1>
F:1534-1634/Domain: collagen IV carboxyl-terminal repeat <CT1>
F:1644-1748/Domain: collagen IV carboxyl-terminal repeat <CT2>
F:129/Modified site: allysine (Lys) #status predicted

Query Match      8.5%; Score 125.5; DB 2; Length 1752;
Best Local Similarity 25.0%; Pred. No. 0.0077;
Matches 38; Conservative 20; Mismatches 51; Indels 43; Gaps 6;

QY      90 LOGPAGVPGRDSPGANGIPGTGIGIRGDFGKGEKGECLRESF----- 132
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB      1471 LITDCKTTPGVQGNPGVSGVGPBGLKGBQCFRGONCPGDPGHPGTGGBAIGIPGSSGTF 1530
QY      133 --EESWTPNYKQC-----SW-----SSINYGIDLGKIAECTFTKMSNSALRYL 174
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB      1531 ITRHSQTTTSPQCPQGTAKMMHGYSLFLVQGNRNGHGDGLGRKQSC-----LKRFSIMPL 1586
QY      175 FSGSLRLKCRNACCGQRYFTFGAECGSLPT 206
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB      1587 FC-NINNVCHVASRNDYSYWLSTTE--PMPM 1614

RESULT 4
CGHUTL
collagen alpha 1(III) chain precursor - human
N:Alternate names: procollagen alpha 1(III) chain
```

```
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence revision 01-Sep-1995 #text change 09-Jul-2004
C:Accession: S05272; S04642; FE0011; S01726; S04887; A50399; A94562; I51868; S59511; A90
R:Prockop, D.J.
submitted to the EMBL Data Library, February 1989
A:Reference number: S05272
A:Accession: S05272
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1240, 'V', 1242-1466 <PRC>
A:Cross-references: UNIPROT:P02461; EMBL:X14420; NID:g30057; PIDN:CAA32583.1; PID:g30058
R:Ala-Kokko, L.; Kontusaari, S.; Baldwin, C.T.; Kuitvanen, H.; Prockop, D.J.
Biochem. J. 260, 509-516, 1989
A:Title: Structure of cDNA clones coding for the entire prepro-alpha1(III) chain of huma
erences.
A:Reference number: S04642; MUID:89350838; PMID:2764886
A:Accession: S04642
A:Molecule type: mRNA
A:Residues: 1-1196 <ALA>
A:Cross-references: EMBL:X14420; NID:g30057; PIDN:CAA32583.1; PID:g30058
A:Note: the complete sequence is not shown
R:Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.
Gene 78, 255-265, 1989
A:Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene (
A:Reference number: PE0011; MUID:89378752; PMID:2777083
A:Accession: PE0011
A:Molecule type: DNA
A:Residues: 1-176 <BEN>
A:Cross-references: GB:M26939; NID:g180813; PIDN:AAA52040.1; PID:g180814
R:Tomar, P.D.; Ricca, G.A.; de Crombrughe, B.
Nucleic Acids Res. 16, 7201, 1988
A:Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human pre
A:Reference number: S01726; MUID:88303360; PMID:3405773
A:Accession: S01726
A:Molecule type: mRNA
A:Residues: 1-170 <TOM>
A:Cross-references: EMBL:X07240; NID:g30060; PIDN:CAA30229.1; PID:g30061
A:Note: the authors translated the codon CAG for residue 154 as His
R:Janeczko, R.A.; Ramirez, F.
Nucleic Acids Res. 17, 6742, 1989
A:Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen.
A:Reference number: S04887; MUID:89386015; PMID:2780304
A:Accession: S04887
A:Molecule type: mRNA
A:Residues: 149-163, 'G', 164-240, 'D', 242-471, 'D', 473-487, 'I', 489, 'S', 491-613, 'Y', 615-634,
A:Cross-references: EMBL:X15332; NID:q29545; PIDN:CAA33387.1; PID:g930045
R:Sever, J.M.; Kang, A.H.
Biochemistry 16, 1158-1164, 1977
A:Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptide
A:Reference number: A50399; MUID:77134724; PMID:557335
A:Accession: A50399
A:Molecule type: Protein
A:Residues: 'V', 169-225, 229-232, 'P', 234-292, 'D', 294-398 <SEY1>
A:Experimental source: liver
A:Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galact
R:Sever, J.M.
submitted to the Atlas, December 1977
A:Reference number: A94562
A:Accession: A94562
A:Molecule type: Protein
A:Residues: 'V', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>
A:Experimental source: liver
A:Note: author submitted corrections to A50399
R:Miliewicz, D.M.; Witz, A.M.; Smith, A.C.; Manchester, D.K.; Waldestein, G.; Byers, P.H.
Am. J. Hum. Genet. 53, 62-70, 1993
A:Title: Parental somatic and germ-line mosaicism for a multiexon deletion with unusual
splicing.
A:Reference number: I51868; MUID:93304430; PMID:8317500
A:Accession: I51868
A:Status: preliminary; translated from GB/EMBL/DDBU
A:Molecule type: DNA
A:Residues: 186-194 <MIL>
```

A:Cross-references: GB:562925; NID:9386425; PIDN:AA013937.1; PID:94261637  
A:Chilod, A.A.; Sillence, D.O.; Cole, W.G.; Bateman, J.F.  
Biochem. J. 311, 939-943, 1995  
A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3A1 gene  
A:Reference number: 559511; PMID:96067614; PMID:7487954  
A:Accession: S59511  
A:Molecule type: mRNA  
A:Residues: 302-423 <CHT>  
A:Cross-references: GB:579877; NID:91195576; PIDN:AA035615.1; PID:91195577  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 17, 3404-3411, 1978  
A:Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBr peptides  
A:Reference number: A90414; PMID:79000343; PMID:687591  
A:Accession: A90414  
A:Molecule type: protein  
A:Residues: 399-675, 'N', 677-727 <SEY3>  
A:Experimental source: liver  
R:Lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.  
J. Biol. Chem. 266, 5256-5259, 1991  
A:Title: G to T transversion at position +5 of a splice donor site causes skipping of the alpha 1(I) collagen gene  
A:Reference number: 155349; PMID:91161621; PMID:1672129  
A:Accession: 155349  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 537-605 <LEB>  
A:Cross-references: GB:M59312; NID:9180815; PIDN:AA052041.1; PID:9180816  
R:Sever, J.M.; Mainardi, C.; Kang, A.H.  
Biochemistry 19, 1583-1589, 1980  
A:Title: Covalent structure of collagen: amino acid sequence of alpha 1(III)-CB5 from type I collagen  
A:Reference number: A90438; PMID:80198282; PMID:6246925  
A:Accession: A90438  
A:Molecule type: protein  
A:Residues: 728-895, 'A', 897-964 <SEY4>  
A:Experimental source: liver  
R:Cole, W.G.; Chiodo, A.A.; Lemande, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Chan J. Biol. Chem. 265, 17070-17077, 1990  
A:Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping in a patient with Ehlers-Danlos syndrome  
A:Reference number: A38303; PMID:91009133; PMID:2145268  
A:Accession: A38303  
A:Molecule type: mRNA  
A:Residues: 861-1015 <COL>  
A:Cross-references: GB:005617; GB:M55603; GB:M59227; NID:9180878; PIDN:AA059383.1; PID:9180878  
A:Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syndrome  
R:Manco, B.S.; Dalgleish, R.  
Nucleic Acids Res. 16, 2337, 1988  
A:Title: Human pro alpha 1(III) collagen: cDNA sequence for the 3' end.  
A:Reference number: S02119; PMID:8819827; PMID:3357782  
A:Accession: S02119  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>  
A:Cross-references: EMBL:X06700; NID:930053; PIDN:CAA29886.1; PID:930054  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 20, 2621-2627, 1981  
A:Title: Covalent structure of collagen: amino acid sequence of alpha 1(III)-CB9 from type I collagen  
A:Reference number: A90446; PMID:81208139; PMID:7016180  
A:Accession: A90446  
A:Molecule type: protein  
A:Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 1157-1158  
A:Experimental source: liver  
R:Gold, H.R.; Brinker, J.M.; May, M.; Pihlajantien, T.; Morrow, S.; Rosenbloom, J.; Myer Nucleic Acids Res. 12, 9383-9394, 1984  
A:Title: Molecular cloning and cDNA sequence of human type III procollagen alpha 1(I) gene  
A:Reference number: A93551; PMID:85087944; PMID:6096827  
A:Accession: A93551  
A:Molecule type: mRNA  
A:Residues: 1065-1155, 'P', 1157-1466 <LOI>  
A:Cross-references: EMBL:X01655; EMBL:X01742; NID:929584; PIDN:CAA25821.1; PID:929584  
R:Sever, J.M.; Dalgleish, R.; Kluge-Beckerman, B.; Remard, S.I.; Tolstoshev, P.; Brant Biochemistry 25, 1408-1413, 1986  
A:Title: Human type III collagen gene expression is coordinately modulated with the type I collagen gene  
A:Reference number: 152393; PMID:86187804; PMID:3754462  
A:Accession: 152393

A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1161-1200 <MIS>  
A:Cross-references: GB:M13146; NID:9180415; PIDN:AA052003.1; PID:9180416  
R:Emmanuel, B.S.; Camizazaro, L.A.; Sever, J.M.; Myer, J.C.  
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985  
A:Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm of chromosome 12  
A:Reference number: 159025; PMID:85216505; PMID:3858826  
A:Accession: 159025  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1165-1196 <EMA>  
A:Cross-references: GB:M11134; NID:9180417; PIDN:AA052004.1; PID:9180418  
R:Chu, M.L.; Wei, D.; de Wet, W.; Bernard, M.; Stippola, M.; Ramirez, F.  
J. Biol. Chem. 260, 4357-4363, 1985  
A:Title: Isolation of cDNA and genomic clones encoding human pro-alpha 1(III) collagen  
A:Reference number: A92516; PMID:85157600; PMID:2579949  
A:Accession: A92516  
A:Molecule type: DNA  
A:Residues: 1176-1240, 'V', 1242-1356, 'P', 1358-1466 <CHU>  
A:Cross-references: GB:M10615; GB:M10793; GB:M10794; GB:M10795; GB:M10796; GB:M10797; GB:M10798  
A:Experimental source: liver  
A:Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given are  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit are  
3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently  
C:Genetics:  
A:Gene: GDB:COL3A1  
A:Cross-references: GDB:118729; OMIM:120180  
A:Map position: 2q31-2q31  
A:Introns: 27/1, 94/3, 111/3, 149/3, 176/3, 554/3, 587/3, 1175/3, 1275/1, 1337/3, 1418/3, 1419/3, 1420/3, 1421/3, 1422/3, 1423/3, 1424/3, 1425/3, 1426/3, 1427/3, 1428/3, 1429/3, 1430/3, 1431/3, 1432/3, 1433/3, 1434/3, 1435/3, 1436/3, 1437/3, 1438/3, 1439/3, 1440/3, 1441/3, 1442/3, 1443/3, 1444/3, 1445/3, 1446/3, 1447/3, 1448/3, 1449/3, 1450/3, 1451/3, 1452/3, 1453/3, 1454/3, 1455/3, 1456/3, 1457/3, 1458/3, 1459/3, 1460/3, 1461/3, 1462/3, 1463/3, 1464/3, 1465/3, 1466/3, 1467/3, 1468/3, 1469/3, 1470/3, 1471/3, 1472/3, 1473/3, 1474/3, 1475/3, 1476/3, 1477/3, 1478/3, 1479/3, 1480/3, 1481/3, 1482/3, 1483/3, 1484/3, 1485/3, 1486/3, 1487/3, 1488/3, 1489/3, 1490/3, 1491/3, 1492/3, 1493/3, 1494/3, 1495/3, 1496/3, 1497/3, 1498/3, 1499/3, 1500/3, 1501/3, 1502/3, 1503/3, 1504/3, 1505/3, 1506/3, 1507/3, 1508/3, 1509/3, 1510/3, 1511/3, 1512/3, 1513/3, 1514/3, 1515/3, 1516/3, 1517/3, 1518/3, 1519/3, 1520/3, 1521/3, 1522/3, 1523/3, 1524/3, 1525/3, 1526/3, 1527/3, 1528/3, 1529/3, 1530/3, 1531/3, 1532/3, 1533/3, 1534/3, 1535/3, 1536/3, 1537/3, 1538/3, 1539/3, 1540/3, 1541/3, 1542/3, 1543/3, 1544/3, 1545/3, 1546/3, 1547/3, 1548/3, 1549/3, 1550/3, 1551/3, 1552/3, 1553/3, 1554/3, 1555/3, 1556/3, 1557/3, 1558/3, 1559/3, 1560/3, 1561/3, 1562/3, 1563/3, 1564/3, 1565/3, 1566/3, 1567/3, 1568/3, 1569/3, 1570/3, 1571/3, 1572/3, 1573/3, 1574/3, 1575/3, 1576/3, 1577/3, 1578/3, 1579/3, 1580/3, 1581/3, 1582/3, 1583/3, 1584/3, 1585/3, 1586/3, 1587/3, 1588/3, 1589/3, 1590/3, 1591/3, 1592/3, 1593/3, 1594/3, 1595/3, 1596/3, 1597/3, 1598/3, 1599/3, 1600/3, 1601/3, 1602/3, 1603/3, 1604/3, 1605/3, 1606/3, 1607/3, 1608/3, 1609/3, 1610/3, 1611/3, 1612/3, 1613/3, 1614/3, 1615/3, 1616/3, 1617/3, 1618/3, 1619/3, 1620/3, 1621/3, 1622/3, 1623/3, 1624/3, 1625/3, 1626/3, 1627/3, 1628/3, 1629/3, 1630/3, 1631/3, 1632/3, 1633/3, 1634/3, 1635/3, 1636/3, 1637/3, 1638/3, 1639/3, 1640/3, 1641/3, 1642/3, 1643/3, 1644/3, 1645/3, 1646/3, 1647/3, 1648/3, 1649/3, 1650/3, 1651/3, 1652/3, 1653/3, 1654/3, 1655/3, 1656/3, 1657/3, 1658/3, 1659/3, 1660/3, 1661/3, 1662/3, 1663/3, 1664/3, 1665/3, 1666/3, 1667/3, 1668/3, 1669/3, 1670/3, 1671/3, 1672/3, 1673/3, 1674/3, 1675/3, 1676/3, 1677/3, 1678/3, 1679/3, 1680/3, 1681/3, 1682/3, 1683/3, 1684/3, 1685/3, 1686/3, 1687/3, 1688/3, 1689/3, 1690/3, 1691/3, 1692/3, 1693/3, 1694/3, 1695/3, 1696/3, 1697/3, 1698/3, 1699/3, 1700/3, 1701/3, 1702/3, 1703/3, 1704/3, 1705/3, 1706/3, 1707/3, 1708/3, 1709/3, 1710/3, 1711/3, 1712/3, 1713/3, 1714/3, 1715/3, 1716/3, 1717/3, 1718/3, 1719/3, 1720/3, 1721/3, 1722/3, 1723/3, 1724/3, 1725/3, 1726/3, 1727/3, 1728/3, 1729/3, 1730/3, 1731/3, 1732/3, 1733/3, 1734/3, 1735/3, 1736/3, 1737/3, 1738/3, 1739/3, 1740/3, 1741/3, 1742/3, 1743/3, 1744/3, 1745/3, 1746/3, 1747/3, 1748/3, 1749/3, 1750/3, 1751/3, 1752/3, 1753/3, 1754/3, 1755/3, 1756/3, 1757/3, 1758/3, 1759/3, 1760/3, 1761/3, 1762/3, 1763/3, 1764/3, 1765/3, 1766/3, 1767/3, 1768/3, 1769/3, 1770/3, 1771/3, 1772/3, 1773/3, 1774/3, 1775/3, 1776/3, 1777/3, 1778/3, 1779/3, 1780/3, 1781/3, 1782/3, 1783/3, 1784/3, 1785/3, 1786/3, 1787/3, 1788/3, 1789/3, 1790/3, 1791/3, 1792/3, 1793/3, 1794/3, 1795/3, 1796/3, 1797/3, 1798/3, 1799/3, 1800/3, 1801/3, 1802/3, 1803/3, 1804/3, 1805/3, 1806/3, 1807/3, 1808/3, 1809/3, 1810/3, 1811/3, 1812/3, 1813/3, 1814/3, 1815/3, 1816/3, 1817/3, 1818/3, 1819/3, 1820/3, 1821/3, 1822/3, 1823/3, 1824/3, 1825/3, 1826/3, 1827/3, 1828/3, 1829/3, 1830/3, 1831/3, 1832/3, 1833/3, 1834/3, 1835/3, 1836/3, 1837/3, 1838/3, 1839/3, 1840/3, 1841/3, 1842/3, 1843/3, 1844/3, 1845/3, 1846/3, 1847/3, 1848/3, 1849/3, 1850/3, 1851/3, 1852/3, 1853/3, 1854/3, 1855/3, 1856/3, 1857/3, 1858/3, 1859/3, 1860/3, 1861/3, 1862/3, 1863/3, 1864/3, 1865/3, 1866/3, 1867/3, 1868/3, 1869/3, 1870/3, 1871/3, 1872/3, 1873/3, 1874/3, 1875/3, 1876/3, 1877/3, 1878/3, 1879/3, 1880/3, 1881/3, 1882/3, 1883/3, 1884/3, 1885/3, 1886/3, 1887/3, 1888/3, 1889/3, 1890/3, 1891/3, 1892/3, 1893/3, 1894/3, 1895/3, 1896/3, 1897/3, 1898/3, 1899/3, 1900/3, 1901/3, 1902/3, 1903/3, 1904/3, 1905/3, 1906/3, 1907/3, 1908/3, 1909/3, 1910/3, 1911/3, 1912/3, 1913/3, 1914/3, 1915/3, 1916/3, 1917/3, 1918/3, 1919/3, 1920/3, 1921/3, 1922/3, 1923/3, 1924/3, 1925/3, 1926/3, 1927/3, 1928/3, 1929/3, 1930/3, 1931/3, 1932/3, 1933/3, 1934/3, 1935/3, 1936/3, 1937/3, 1938/3, 1939/3, 1940/3, 1941/3, 1942/3, 1943/3, 1944/3, 1945/3, 1946/3, 1947/3, 1948/3, 1949/3, 1950/3, 1951/3, 1952/3, 1953/3, 1954/3, 1955/3, 1956/3, 1957/3, 1958/3, 1959/3, 1960/3, 1961/3, 1962/3, 1963/3, 1964/3, 1965/3, 1966/3, 1967/3, 1968/3, 1969/3, 1970/3, 1971/3, 1972/3, 1973/3, 1974/3, 1975/3, 1976/3, 1977/3, 1978/3, 1979/3, 1980/3, 1981/3, 1982/3, 1983/3, 1984/3, 1985/3, 1986/3, 1987/3, 1988/3, 1989/3, 1990/3, 1991/3, 1992/3, 1993/3, 1994/3, 1995/3, 1996/3, 1997/3, 1998/3, 1999/3, 2000/3, 2001/3, 2002/3, 2003/3, 2004/3, 2005/3, 2006/3, 2007/3, 2008/3, 2009/3, 2010/3, 2011/3, 2012/3, 2013/3, 2014/3, 2015/3, 2016/3, 2017/3, 2018/3, 2019/3, 2020/3, 2021/3, 2022/3, 2023/3, 2024/3, 2025/3, 2026/3, 2027/3, 2028/3, 2029/3, 2030/3, 2031/3, 2032/3, 2033/3, 2034/3, 2035/3, 2036/3, 2037/3, 2038/3, 2039/3, 2040/3, 2041/3, 2042/3, 2043/3, 2044/3, 2045/3, 2046/3, 2047/3, 2048/3, 2049/3, 2050/3, 2051/3, 2052/3, 2053/3, 2054/3, 2055/3, 2056/3, 2057/3, 2058/3, 2059/3, 2060/3, 2061/3, 2062/3, 2063/3, 2064/3, 2065/3, 2066/3, 2067/3, 2068/3, 2069/3, 2070/3, 2071/3, 2072/3, 2073/3, 2074/3, 2075/3, 2076/3, 2077/3, 2078/3, 2079/3, 2080/3, 2081/3, 2082/3, 2083/3, 2084/3, 2085/3, 2086/3, 2087/3, 2088/3, 2089/3, 2090/3, 2091/3, 2092/3, 2093/3, 2094/3, 2095/3, 2096/3, 2097/3, 2098/3, 2099/3, 2100/3, 2101/3, 2102/3, 2103/3, 2104/3, 2105/3, 2106/3, 2107/3, 2108/3, 2109/3, 2110/3, 2111/3, 2112/3, 2113/3, 2114/3, 2115/3, 2116/3, 2117/3, 2118/3, 2119/3, 2120/3, 2121/3, 2122/3, 2123/3, 2124/3, 2125/3, 2126/3, 2127/3, 2128/3, 2129/3, 2130/3, 2131/3, 2132/3, 2133/3, 2134/3, 2135/3, 2136/3, 2137/3, 2138/3, 2139/3, 2140/3, 2141/3, 2142/3, 2143/3, 2144/3, 2145/3, 2146/3, 2147/3, 2148/3, 2149/3, 2150/3, 2151/3, 2152/3, 2153/3, 2154/3, 2155/3, 2156/3, 2157/3, 2158/3, 2159/3, 2160/3, 2161/3, 2162/3, 2163/3, 2164/3, 2165/3, 2166/3, 2167/3, 2168/3, 2169/3, 2170/3, 2171/3, 2172/3, 2173/3, 2174/3, 2175/3, 2176/3, 2177/3, 2178/3, 2179/3, 2180/3, 2181/3, 2182/3, 2183/3, 2184/3, 2185/3, 2186/3, 2187/3, 2188/3, 2189/3, 2190/3, 2191/3, 2192/3, 2193/3, 2194/3, 2195/3, 2196/3, 2197/3, 2198/3, 2199/3, 2200/3, 2201/3, 2202/3, 2203/3, 2204/3, 2205/3, 2206/3, 2207/3, 2208/3, 2209/3, 2210/3, 2211/3, 2212/3, 2213/3, 2214/3, 2215/3, 2216/3, 2217/3, 2218/3, 2219/3, 2220/3, 2221/3, 2222/3, 2223/3, 2224/3, 2225/3, 2226/3, 2227/3, 2228/3, 2229/3, 2230/3, 2231/3, 2232/3, 2233/3, 2234/3, 2235/3, 2236/3, 2237/3, 2238/3, 2239/3, 2240/3, 2241/3, 2242/3, 2243/3, 2244/3, 2245/3, 2246/3, 2247/3, 2248/3, 2249/3, 2250/3, 2251/3, 2252/3, 2253/3, 2254/3, 2255/3, 2256/3, 2257/3, 2258/3, 2259/3, 2260/3, 2261/3, 2262/3, 2263/3, 2264/3, 2265/3, 2266/3, 2267/3, 2268/3, 2269/3, 2270/3, 2271/3, 2272/3, 2273/3, 2274/3, 2275/3, 2276/3, 2277/3, 2278/3, 2279/3, 2280/3, 2281/3, 2282/3, 2283/3, 2284/3, 2285/3, 2286/3, 2287/3, 2288/3, 2289/3, 2290/3, 2291/3, 2292/3, 2293/3, 2294/3, 2295/3, 2296/3, 2297/3, 2298/3, 2299/3, 2300/3, 2301/3, 2302/3, 2303/3, 2304/3, 2305/3, 2306/3, 2307/3, 2308/3, 2309/3, 2310/3, 2311/3, 2312/3, 2313/3, 2314/3, 2315/3, 2316/3, 2317/3, 2318/3, 2319/3, 2320/3, 2321/3, 2322/3, 2323/3, 2324/3, 2325/3, 2326/3, 2327/3, 2328/3, 2329/3, 2330/3, 2331/3, 2332/3, 2333/3, 2334/3, 2335/3, 2336/3, 2337/3, 2338/3, 2339/3, 2340/3, 2341/3, 2342/3, 2343/3, 2344/3, 2345/3, 2346/3, 2347/3, 2348/3, 2349/3, 2350/3, 2351/3, 2352/3, 2353/3, 2354/3, 2355/3, 2356/3, 2357/3, 2358/3, 2359/3, 2360/3, 2361/3, 2362/3, 2363/3, 2364/3, 2365/3, 2366/3, 2367/3, 2368/3, 2369/3, 2370/3, 2371/3, 2372/3, 2373/3, 2374/3, 2375/3, 2376/3, 2377/3, 2378/3, 2379/3, 2380/3, 2381/3, 2382/3, 2383/3, 2384/3, 2385/3, 2386/3, 2387/3, 2388/3, 2389/3, 2390/3, 2391/3, 2392/3, 2393/3, 2394/3, 2395/3, 2396/3, 2397/3, 2398/3, 2399/3, 2400/3, 2401/3, 2402/3, 2403/3, 2404/3, 2405/3, 2406/3, 2407/3, 2408/3, 2409/3, 2410/3, 2411/3, 2412/3, 2413/3, 2414/3, 2415/3, 2416/3, 2417/3, 2418/3, 2419/3, 2420/3, 2421/3, 2422/3, 2423/3, 2424/3, 2425/3, 2426/3, 2427/3, 2428/3, 2429/3, 2430/3, 2431/3, 2432/3, 2433/3, 2434/3, 2435/3, 2436/3, 2437/3, 2438/3, 2439/3, 2440/3, 2441/3, 2442/3, 2443/3, 2444/3, 2445/3, 2446/3, 2447/3, 2448/3, 2449/3, 2450/3, 2451/3, 2452/3, 2453/3, 2454/3, 2455/3, 2456/3, 2457/3, 2458/3, 2459/3, 2460/3, 2461/3, 2462/3, 2463/3, 2464/3, 2465/3, 2466/3, 2467/3, 2468/3, 2469/3, 2470/3, 2471/3, 2472/3, 2473/3, 2474/3, 2475/3, 2476/3, 2477/3, 2478/3, 2479/3, 2480/3, 2481/3, 2482/3, 2483/3, 2484/3, 2485/3, 2486/3, 2487/3, 2488/3, 2489/3, 2490/3, 2491/3, 2492/3, 2493/3, 2494/3, 2495/3, 2496/3, 2497/3, 2498/3, 2499/3, 2500/3, 2501/3, 2502/3, 2503/3, 2504/3, 2505/3, 2506/3, 2507/3, 2508/3, 2509/3, 2510/3, 2511/3, 2512/3, 2513/3, 2514/3, 2515/3, 2516/3, 2517/3, 2518/3, 2519/3, 2520/3, 2521/3, 2522/3, 2523/3, 2524/3, 2525/3, 2526/3, 2527/3, 2528/3, 2529/3, 2530/3, 2531/3, 2532/3, 2533/3, 2534/3, 2535/3, 2536/3, 2537/3, 2538/3, 2539/3, 2540/3, 2541/3, 2542/3, 2543/3, 2544/3, 2545/3, 2546/3, 2547/3, 2548/3, 2549/3, 2550/3, 2551/3, 2552/3, 2553/3, 2554/3, 2555/3, 2556/3, 2557/3, 2558/3, 2559/3, 2560/3, 2561/3, 2562/3, 2563/3, 2564/3, 2565/3, 2566/3, 2567/3, 2568/3, 2569/3, 2570/3, 2571/3, 2572/3, 2573/3, 2574/3, 2575/3, 2576/3, 2577/3, 2578/3, 2579/3, 2580/3, 2581/3, 2582/3, 2583/3, 2584/3, 2585/3, 2586/3, 2587/3, 2588/3, 2589/3, 2590/3, 2591/3, 2592/3, 2593/3, 2594/3, 2595/3, 2596/3, 2597/3, 2598/

```
RESULT 5
T15268
hypothetical protein F59E12.12 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15268
R:Johnson, D.
Submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid F59E12.
A:Reference number: Z18318
A:Accession: T15268
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-311 <OH>
A:Cross-references: UNIPROT:O01904; EMBL:AF003386; NID:G2088833; PID:G2088834; PIDN:AA85
A:Experimental source: strain Bristol N2; clone F59E12
C:Genetics:
A:Gene: CESP:F59E12.12
A:Map position: 2
A:introns: 24/2

Query Match      8.4%; Score 123.5; DB 2; Length 311;
Best Local Similarity 53.3%; Pred. No. 0.0016;
Matches 24; Conservative 4; Mismatches 12; Indels 5; Gaps 1;

QY 87 GNC-----LQGPAGVGRDSSPGANGIPGTGIPGRDGFKEGKE 126
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 227 GKCDEVNVVAQGPSPGPPGLPGPDGLPGTPGNPGQDGRGPAGE 271

RESULT 6
T29731
hypothetical protein F41F3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29731
R:Badcock, K.; Le, T.T.; Gattung, S.
Submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid F41F3.
A:Reference number: Z20674
A:Accession: T29731
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-310 <BAD>
A:Cross-references: UNIPROT:Q20822; EMBL:U55366; PIDN:AAA97982.1; GSPDB:GN00023; CESP:F4
A:Experimental source: strain Bristol N2; clone F41F3
C:Genetics:
A:Gene: CESP:F41F3.4
A:Map position: 5
A:introns: 27/3; 68/3

Query Match      8.3%; Score 122.5; DB 2; Length 310;
Best Local Similarity 24.4%; Pred. No. 0.002;
Matches 39; Conservative 9; Mismatches 39; Indels 73; Gaps 5;

QY 4 AAEIGTLSBAGICRPLCLILCASQPLPHMPPGGRASPOLLIGLFLVLLILLIQLSAPSS 63
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 93 AVEVASRVSGSCQCCCL-----PGPPGTPPGPGRPKP-----GRPGA 132

QY 64 ASEN-----PKYKQKALIRQREVVVLVNGMCLQGPAGVPGRDSPGANG 107
    |||||:|||||:|||||:|||||:|||||:|||||:
DB 133 CGNHGMPGKPTGLPCDPVTVTPCKP-----PPGPAEPGHGAPGAPG 176

QY 108 IPGTGPI-----PGRDGFKEGKE 126
    |||||:|||||:|||||:|||||:
DB 177 KEGAPGVGGSSGAPGAPGPKGAPGPGQGRDQPGQAQ 216

RESULT 7
A45748
collagen alpha 1(VII) chain - mouse (fragment)
```

```
C:Species: Mus musculus (house mouse)
C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A45748
R:Li, K.; Christiano, A.M.; Copeland, N.G.; Gilbert, D.J.; Chu, M.L.; Jenkins, N.A.; Ut
Genomics 16, 733-739, 1993
A:Title: CDNA cloning and chromosomal mapping of the mouse type VII collagen gene (Col7a
A:Reference number: A45748; MUID:93315168; PMID:832648
A:Accession: A45748
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-920 <LII>
A:Cross-references: UNIPROT:O63870; GB:S63654; NID:9386656; PIDN:AA827492.1; PID:9386657
A:Experimental source: epidermal keratinocyte
A>Note: sequence extracted from NCBI backbone (NCBIN:135000, NCBIPI:135001)

Query Match      8.3%; Score 122.5; DB 2; Length 920;
Best Local Similarity 63.2%; Pred. No. 0.0069;
Matches 24; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 92 GPAGVPR---DGSRGANGIPGTGIPGRDGFKEGKE 126
    |||||:|||||:|||||:|||||:|||||:|||||:
DB 680 GPPVPGKAGDDGKPGKGLNGKNGPDGPDGDKGKEGD 717

RESULT 8
I48103
type VII collagen - Chinese hamster (fragment)
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I48103
R:Greenspan, D.S.
Hum. Mol. Genet. 2, 273-278, 1993
A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
A:Reference number: I48103; MUID:93771985; PMID:8499916
A:Accession: I48103
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1549 <RGS>
A:Cross-references: UNIPROT:O60444; GB:I06863; NID:9388624; PIDN:AAA36968.1; PID:9388625
F:I484-1336/Domain: animal Kunitz-type proteinase inhibitor homology <BRI>

Query Match      8.3%; Score 122.5; DB 2; Length 1549;
Best Local Similarity 63.2%; Pred. No. 0.012;
Matches 24; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 92 GPAGVPR---DGSRGANGIPGTGIPGRDGFKEGKE 126
    |||||:|||||:|||||:|||||:|||||:|||||:
DB 426 GPPVPGKAGDDGKPGKGLNGKNGPDGPDGDKGKEGD 463

RESULT 9
A39762
collagen alpha 1(XIV) chain - bovine (fragments)
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Feb-1992 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C:Accession: A39762; S23815; S16905
R:Hubert-Foucher, B.; Font, B.; Eichenberger, D.; Goldschmidt, D.; Lethias, C.; van der
J. Biol. Chem. 267, 15759-15764, 1992
A:Title: Purification and characterization of native type XIV collagen.
A:Reference number: S23815; MUID:92348439; PMID:1322405
A:Accession: S23815
A:Molecule type: protein
A:Residues: 'XR', 82-94, 'X', 96 <AUB>
C:Complex: homotrimer
C:Keywords: coiled coil; disulfide bond; extracellular matrix; homotrimer; hydroxylysine
```

```

Query Match 8.3% Score 122; DB 2; Length 178;
Best Local Similarity 32.7%; Pred. No.0.0012;
Matches 32; Conservative 8; Mismatches 34; Indels 24; Gaps 4;

QY 32 TWHPGRAASPOLLLGLFLVLLLLQLLSAPSASSENPKVKOKALIRPQREVVDLYNQCICQ 91
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 46 SMGPFGALGPP-----GVPFAGD-----LQSQAVNRQ-VCEQLQSHRIQ 84
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 92 GPAGVPGDGSFGA---NGIPQTGIPGRDGEKGEKGE 126
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 85 GPPGEPGRPGSGPTGEGQPPGAPGPPGNAGVPGSGXE 122
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
collagen alpha 1(XIV) chain - human (fragments)
C|Species: Homo sapiens (man)
C|Date: 15-Jul-1995 #sequence revision 03-Nov-1995 #text_change 09-Jul-2004
C|Accession: S46657; S37750; A44398
R|Brown, J.C.; Golbik, R.; Mann, K.; Timpl, R.
Matrix Biol. 14, 287-295, 1994
A|Title: Structure and stability of the triple-helical domains of human collagen XIV.
A|Reference number: S46657; MUID:95128518; PMID:7827751
A|Accession: S46657
A|Status: preliminary
A|Molecule type: protein
A|Residues: 26-330 <BRO>
A|Cross-references: UNIPROT:Q9UDF6
R|Brown, J.C.; Golbik, R.; Mann, K.; Timpl, R.
Submitted to the EMBL Data Library, October 1993
A|Description: Structure and stability of the triple helical domains of human collagen X
A|Reference number: S37749
A|Accession: S37750
A|Molecule type: protein
A|Residues: 203-330 <BR2>
R|Brown, J.C.; Mann, K.; Wiedemann, H.; Timpl, R.
J. Cell Biol. 120, 557-567, 1993
A|Title: Structure and binding properties of collagen type XIV isolated from human placenta
A|Reference number: A44398; MUID:95132085; PMID:8421066
A|Accession: A44398
A|Molecule type: protein
A|Residues: 1-17;18-25;26-42,'X',44-45;74-84,100-117;136-150,'SM';238-260,'X',262-263,'X'
A|Experimental source: placenta
C|Comment: Prolines and lysines at the third position of the tripeptide repeating unit (ed and subsequently O-glycosylated).
C|Genetics:
A|Gene: COL14A1
C|Complex: type XIV collagen may be a heterotrimer of alpha 1(XIV) chains and one alpha 1(XIV) chain.
C|Function:
A|Description: structural component of extracellular fibrous polymer associated with collagen
C|Keywords: collod coll, extracellular matrix; glycoprotein; hydroxylysine; hydroxyproline
F|374,84,111,117,141,144,147,219,222,225,228,231,237,240,243,249,252,267,278,292,295,298,
F|374,206,288/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F|43,261/Modified site: 5-hydroxylysine (Lys) #status experimental
F|43,261,264/Binding site: carboxyrate (Lys) (covalent) #status experimental
F|264/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental

Query Match 8.3%; Score 122; DB 2; Length 330;
Best Local Similarity 31.9%; Pred. No.0.0024;
Matches 36; Conservative 10; Mismatches 33; Indels 34; Gaps 5;

QY 35 PGRGAASPOLLLGLFLVLLLLQLLSAPSASSENPKVKOKALIR-OREVD-----L 84
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 155 POGALGPPGV-----PGAKGERGERDGLSQAVNRKVAQAQVBLQISHMAR 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 85 YNGM-----CLOGPAGVPGDGSFGA---NGIPQTGIPGRDGEKGEKGE 126
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 202 YNOIPSHSSIRTVQGPGEGRPGSGPCAPGEGQPPGPPGPPGNAGVPGTPGE 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

[illegible]

C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Jul-2004

C:Accession: A61396

R:Citizael, P.E.; Hosoi, J.; Montgomery, J.C.; Wiseman, R.W.; Barrett, J.C.

Mol. Cell. Biochem. 4, 14-24, 1991

A:Title: Loss of a tumor suppressor gene function is correlated with downregulation of

A:Reference number: A61396; MUID:91182265; PMID:2009131

A:Accession: A61396

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-323 <C1>

A:Cross-references: UNIPROT:Q7M099

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology

F:120-323/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 8.1% Score 119.5; DB 2; Length 323;

Best Local Similarity 20.6%; Pred. No.0.004; Mismatches 47; Indels 153; Gaps 10;

Matches 57; Conservative 20; Mismatches 47; Indels 153; Gaps 10;

QY 90 LOGPAGVGRDQSPGAN-----GIRGTP----- 112

DB 9 LTGPAGRGREGSPADGPRGRDGAAGVGRDGTGALGARQAPGAPGAPGATKQKG 68

QY 113 -----GIRGRDGRKGRGE-----CIRSEFE 133

DB 69 DRGSEAGQPMGSPSGPAGARGLPGQGRGKSGSEGEBSLRQHDVEVDATLKSINNQIE 128

QY 134 ESWTP-----NYKOC--SMSSLNYGIDGKFAECTFTPMRS----- 167

DB 129 SIRSPDGSCKPARPCQDLKLCHEPKSGDVIWDNQ--GCTLDAMKAFCMNETGECTVY 186

QY 168 -NSA-----LRVLF--SGSLRLKRNACCORNYFTFNGAEC 200

DB 187 PNPASVPKKNWMSKGEKKNHIVQWTFRLRLSTEGSQNITVHCNK----- 231

QY 201 SGPPLTEALIVYDQSPPELNTINIRHTSSVEGLCEG 237

DB 232 -----SLAVDEAAGNLKKALLIQSSNDVEMRAEG 261

RESULT 14

CGHUB

collagen alpha 3(IV) chain precursor, long splice form - human

N:Alternate names: Goodpasture antigen, procollagen alpha 3(IV) chain long splice form

C:Species: Homo sapiens (man)

C:Date: 28-Oct-1994 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004

A:Accession: A54763; A43928; A44043; A45971; A39786

R:Maruyama, M.; Lehenen, A.; Mochizuki, T.; Tryggvason, K.; Reeder, S.T.

J. Biol. Chem. 269, 23013-23017, 1994

A:Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression

A:Reference number: A54763; MUID:94364994; PMID:8083201

A:Accession: A54763

A:Molecule type: mRNA

A:Residues: 1-1670 <MR>

A:Cross-references: UNIPROT:O01955; GB:X80031; NID:9577563; PID:9577564

R:Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.

J. Clin. Invest. 89, 592-601, 1992

A:Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha

A:Reference number: A43928; MUID:92147878; PMID:1737849

A:Accession: A43928

A:Molecule type: mRNA

A:Residues: 1331-1524, 'I',1526-1670 <TUR>

A:Cross-references: GB:M81379

A:Experimental source: kidney

R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.

J. Biol. Chem. 267, 19780-19784, 1992

A:Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture

A:Reference number: A44043; MUID:93015826; PMID:1400291

A:Accession: A44043

A:Molecule type: DNA; mRNA

A:Residues: 1386-1670 <QUR>

A:Cross-references: GB:M92993; NID:9177895; PID:AAA21610.1; PID:9177896

A>Note: sequence extracted from NCBI backbone (NCBIP:115597)  
R.Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.  
J. Biol. Chem. 269, 17358, 1994  
A:Reference number: A44738; MUID:94274734; PMID:8006044  
A:Contents: annotation; extrim; correction to intronic sequence in A44043  
R.Bernal, D.; Quinones, S.; Saus, J.  
J. Biol. Chem. 269, 12090-12094, 1993  
A>Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.  
A:Accession number: A45971; MUID:93280184; PMID:8505332  
A:Accession: A45971  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1427-1444 <EBR>  
A>Note: sequence extracted from NCBI backbone (NCBIP:133363); sequence incorrectly ident  
R.Morrison, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reeder, S.T.  
Am. J. Hum. Genet. 49, 545-554, 1991  
A>Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of  
A:Reference number: A39786; MUID:91353570; PMID:1682840  
A:Accession: A39786  
A:Molecule type: mRNA  
A:Residues: 1453-1593, 'A', 1595-1670 <MOR>  
A:Cross-references: GB:555790; NID:g234418; PID:AAB19637.1; PID:g234419  
C:Comment: Prolines and lysine at the third position of the tripeptide repeating unit (  
ed and subsequently O-glycosylated.  
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope  
C:Genetics:  
A:Gene: GDB:COL4A3  
A:Cross-references: GDB:128351; OMIM:120070  
A:Map position: 2q36-q37  
A:Introns: 1385/1, 1418/1, 1486/1, 1547/2, 1585/3, 1632/2 #status incomplete  
A>Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands with  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 3(  
mong trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a  
er associations in the interrupted helical domain (with disulfide and desmosine cross-lin  
C:Function:  
A:Description: minor structural component of extracellular basement membrane in kidney g  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracel  
F:1-8/Domain: signal sequence #status predicted <SIG>  
F:29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MAT>  
F:29-42/Domain: amino-terminal nonhelical, NH1 <NH1>  
F:43-1438/Region: interrupted helical  
F:791-793/Region: cell attachment (R-G-D) motif  
F:998-998/Region: cell attachment (R-G-D) motif  
F:1154-1156/Region: cell attachment (R-G-D) motif  
F:1306-1308/Region: cell attachment (R-G-D) motif  
F:1345-1347/Region: cell attachment (R-G-D) motif  
F:1432-1434/Region: cell attachment (R-G-D) motif  
F:1439-1670/Domain: carboxyl-terminal nonhelical, NC1 <NC1>  
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CTR>  
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CTR>  
F:311,33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status predi  
F:1450-Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:1250-1548,1493-Disulfide bonds: (or 1460-1511, 1493-1548) #status predicted  
F:1505-1511,1616-1622/Disulfide bonds: #status predicted  
F:1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted

Query Match	Similarity	Score	DB 1;	Length
Best Local	5.1%;	Pred. No. 0.025;		1670;
Matches	24; Conservative	3; Mismatches	12;	Indels 3; Gaps 1;

91 QGPAGVPGRDSPGANGIPGTPIGRPD---GFKGKGECKLR 129  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 504 QGANGLKSGPSGSFGNTGTLPFPGRPAQSDPGLKGGKGRITLQ 545

RESULT 15  
AS5849  
collagen alpha 1(VII) chain precursor - human  
N:Alternate names: procollagen alpha 1(VII) chain  
C:Species: Homo sapiens (man)  
C:Date: 04-Nov-1994 #sequence\_revision 04-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A54849; PI0844; SI6316; I56328;AA0296; I84686



R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.  
 J. Biol. Chem. 269, 20256-20262, 1994  
 A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(VII)  
 A:Reference number: A54849; PMID:94327588; PMID:8051117  
 A:Accession: A54849  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-2944 <CHR>  
 A:Cross-references: UNIPROT:Q02388; GB:L02870; NID:g987124; PIDN:AAA75438.1; PID:g987125  
 R:Tanaka, T.; Takahashi, K.; Furukawa, F.; Imanura, S.  
 Biochem. Biophys. Res. Commun. 183, 958-963, 1992  
 A:Title: Molecular cloning and characterization of type VII collagen cDNA.  
 A:Reference number: PH0844; PMID:92231902; PMID:1567409  
 A:Accession: PH0844  
 A:Molecule type: mRNA  
 A:Residues: EFR, 340-475, 'RAUSTASHSLCWRATRWHPDNCNRSHTTRACPCPNPASHRAARAG', 524-528, 'C',  
 A:Cross-references: DDBJ:D11152; DDBJ:D13694; NID:g453698; PIDN:BA02853.1; PID:g453699  
 A:Experimental source: Keratinocyte  
 A:Note: the authors translated the codon ACC for residues 394 and 397 as Tyr  
 R:Parente, M.G.; Chung, L.C.; Rymaenen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mat  
 Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991  
 A:Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.  
 A:Reference number: S16316; PMID:91334380; PMID:1871109  
 A:Accession: S16316  
 A:Molecule type: mRNA  
 A:Residues: 815-892, 'E', 894-1439 <PAR>  
 A:Cross-references: GB:M65158; GB:S45017; NID:g180914; PIDN:AAA96439.1; PID:g180915  
 A:Experimental source: Keratinocyte  
 R:Dammann, W.R.; Abernethy, M.L.; Padilla, K.M.; Prisyamb, P.S.; Cook, M.E.; Wright, J.  
 J. Invest. Dermatol. 99, 691-696, 1992  
 A:Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot  
 A:Reference number: I56328; PMID:93107742; PMID:1469284  
 A:Accession: I56328  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 'EFR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>  
 A:Cross-references: GB:S51336; NID:g262308; PIDN:AAB24637.1; PID:g262309  
 R:Seltzer, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.  
 J. Biol. Chem. 264, 3822-3826, 1989  
 A:Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagenase  
 A:Reference number: A30296; PMID:8919437; PMID:2537292  
 A:Accession: A30296  
 A:Molecule type: protein  
 A:Residues: 'A', 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'E', 2032, 'C', 2034-2041, '  
 A:Note: two reported peptides cannot be reliably located  
 R:Greenspan, D.S.  
 Hum. Mol. Genet. 2, 273-278, 1993  
 A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous  
 A:Reference number: I48103; PMID:93271985; PMID:8499916  
 A:Accession: I48103  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 2395-2871, 'S', 2873-2944 <RE2>  
 A:Cross-references: GB:L06862; NID:g388713; PIDN:AAA69196.1; PID:g388714  
 R:Christiano, A.M.; Rymaenen, M.; Uitto, J.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994  
 A:Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser sub  
 A:Reference number: A55255; PMID:94224777; PMID:8170945  
 A:Contents: annotation  
 A:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 ed and subsequently O-glycosylated.  
 C:Genetics:  
 A:Gene: GDB:COL7A1, EBR1, EBD1, EB  
 A:Cross-references: GDB:128750; OMIM:120120  
 A:Map position: 3p21.3-3p21.3  
 A:Note: defects in this gene can result in dominant and recessive dystrophic epidermolys  
 A:Note: there are 118 introns  
 C:Complex: type VII collagen is probably a homotrimer  
 C:Function:  
 A:Description: structural component of extracellular polymer associated with anchoring F  
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli  
 F:1-16/Domain: signal sequence #status predicted <Sig>  
 F:17-2944/Product: collagen alpha 1(VII) chain #status predicted <Mat>

F:17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>  
 F:36-201/Domain: von Willebrand factor type A repeat homology <VWA1>  
 F:231-318/Domain: fibronectin type III repeat homology <FN1>  
 F:327-413/Domain: fibronectin type III repeat homology <FN2>  
 F:414-502/Domain: fibronectin type III repeat homology <FN3>  
 F:508-593/Domain: fibronectin type III repeat homology <FN4>  
 F:598-683/Domain: fibronectin type III repeat homology <FN5>  
 F:686-771/Domain: fibronectin type III repeat homology <FN6>  
 F:776-862/Domain: fibronectin type III repeat homology <FN7>  
 F:864-952/Domain: fibronectin type III repeat homology <FN8>  
 F:954-1045/Domain: fibronectin type III repeat homology <FN9>  
 F:1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>  
 F:1170-1172/Domain: cell attachment (R-G-D) motif  
 F:1189-1253/Domain: cysteine/proline-rich  
 F:1254-1283/Region: interrupted helical  
 F:1334-1336/Region: cell attachment (R-G-D) motif  
 F:2008-2010/Region: cell attachment (R-G-D) motif  
 F:2553-2555/Region: cell attachment (R-G-D) motif  
 F:2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>  
 F:2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
 F:337-786/1109/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:2167-2176, 2185, 2188, 2664, 2667, 2673/Modified site: 4-hydroxyproline (Pro) #status expe  
 F:2625, 2631/Modified site: 5-hydroxylysine (Lys) #status experimental  
 F:2625, 2631/Binding site: carbohydrate (Lys) (covalent) #status experimental  
 F:2634, 2802, 2804/Disulfide bonds: interchain #status predicted

Query Match 8.1%; Score 119.5; DB 2; Length 2944;  
 Best Local Similarity 60.5%; Pred. No. 0.048;  
 Matches 23; Conservative 3; Mismatches 9; Indels 3; Gaps 1;  
 QY 92 GPAGVGR---DGPAGANGPCTPGPGDGRKGE 126  
 DB 1824 GPPGLPKPGEDEKPELNGKNGPDPGDEGRKGEKD 1861

Search completed: August 4, 2005, 04:22:30  
 Job time : 23.0876 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:07:24 ; Search time 89.3529 Seconds  
(without alignments)  
1404.088 Million cell updates/sec

Title: US-10-634-108-2

Perfect score: 1307

Sequence: 1 MHQGRASPOLLLGLFLV.....GDASTGMNSVSRILIELPK 245

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1307	100.0	245	2	08CG08
2	1292	98.9	245	2	08DID6
3	1225	93.7	243	2	096CG8
4	1216	93.0	243	2	06UW91
5	1063	81.3	232	2	081X63
6	932	71.3	231	2	06AXL0
7	133	10.2	287	2	07ZSL3
8	133	10.2	287	2	08CFR0
9	132.5	10.1	565	2	08K036
10	130.5	10.0	551	2	06ZM13
11	129.5	9.9	1669	2	09QZS0
12	128.5	9.8	289	2	08BVD7
13	127.5	9.8	289	2	018799
14	125.5	9.6	291	2	09NAR3
15	125.5	9.6	298	2	018286
16	125.5	9.6	1752	2	007285
17	125	9.6	305	2	09UDP6
18	125	9.6	717	2	09NOS2
19	124.5	9.5	240	2	06LA35
20	124.5	9.5	326	2	0677V9
21	124.5	9.5	358	2	06MEY7
22	124.5	9.5	1336	2	06RZ41
23	124.5	9.5	1516	1	CA1H_HUMAN
24	124.5	9.5	1516	2	06RZ39
25	124.5	9.5	1751	2	06RZ40
26	123.5	9.4	311	2	001904
27	123.5	9.4	312	2	08CHX9
28	123.5	9.4	540	2	086Y22
29	123.5	9.4	957	2	096P44
30	123.5	9.4	957	2	09H0V3
31	123	9.4	1472	2	090ZAO

32	122.5	9.4	282	2	08WP36	08WP36 suberices d
33	122.5	9.4	546	2	06PTU1	06PTU1 mus musculus
34	122.5	9.4	920	2	07BEC6	07BEC6 mus sp. typ
35	122.5	9.4	1549	2	060444	060444 cricetus
36	122.5	9.4	2944	2	063870	063870 mus musculus
37	122	9.3	747	1	CA12_BOVIN	P02459 bos taurus
38	121	9.3	178	2	07MZW9	07MZW9 bos taurus
39	121	9.3	289	1	CQT7_HUMAN	Q9BXJ2 homo sapien
40	121	9.3	1163	2	08N6U4	Q8NEU4 homo sapien
41	121	9.3	1466	1	CA13_HUMAN	P02461 homo sapien
42	121	9.3	1763	1	CA24_ASCSU	P27393 ascaris suu
43	120	9.2	568	2	08CD80	08CD80 mus musculus
44	120	9.2	585	2	080V57	080V57 mus musculus
45	120	9.2	739	2	070575	070575 mus musculus

## ALIGNMENTS

### RESULT 1

ID	Accession	PRELIMINARY	PRT	245 AA.
AC	08CG08			
DT	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Collagen triple helix repeat-containing protein 1.			
GN	Name=Ctrhl1			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley; TISSUE=8 day balloon-injured carotid artery;			
RA	Lehnert W., Moore D.P., Harmon K.J., Mancini M.L., Lindner V.;			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY136824; AAN15748.1; -			
DR	GO: GO:0005737; Cytoplasm; IEA.			
DR	GO: GO:000617; P:phosphate transport; IEA.			
DR	InterPro: IPR008161; Clg_helix.			
DR	InterPro: IPR008160; Collagen.			
DR	Pfam: PF01391; Collagen; 1.			
DR	ProDom: PD000007; Clg_helix; 1.			
KW	Collagen.			
SQ	SEQUENCE 245 AA; 26424 MW; 2296FD6DCDBA21F2 CRC64;			

Query Match 100.0%; Score 1307; DB 2; Length 245;  
Best Local Similarity 100.0%; Pred. No. 4,6e-108;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MHPQGRASPOLLLGLFLVLLLLQLISAPSSASENPKVKALIRQREVVLDLYNGKCLQG	60
DB	1	MHPQGRASPOLLLGLFLVLLLLQLISAPSSASENPKVKALIRQREVVLDLYNGKCLQG	60
QY	61	PAGVPRDRDSPANGIPGTPGIPGRDGFGECECLRESFESSWTPNYQCSSSLNYGI	120
DB	61	PAGVPRDRDSPANGIPGTPGIPGRDGFGECECLRESFESSWTPNYQCSSSLNYGI	120
QY	121	DIGAKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQRYFTFNGACSGPLIEAIIYL	180
DB	121	DIGAKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQRYFTFNGACSGPLIEAIIYL	180
QY	181	DQSPPELNTSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRILII	240
DB	181	DQSPPELNTSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMNSVSRILII	240
QY	241	EEELPK 245	
DB	241	EEELPK 245	

RESULT 2

Q9D1D6 PRELIMINARY; PRT; 245 AA.  
ID Q9D1D6  
AC Q9D1D6;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DR 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched  
DE library, clone:1110014B07 product:hypothetical collagen triple helix  
DE repeat containing protein, full insert sequence.  
GN Name=Ctrnc1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA The FANTOM Consortium.  
RT "The RIKEN Genome Exploration Research Group Phase I & II Team:  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Kono H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arawaka T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai K., Sakai K.,  
RA Sano H., Saeki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AK003674; BAB22930.1; -.  
DR MGD; MGI:1915638; Ctrnc1.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR008161; Clg helix.  
DR InterPro; IPR008160; Collagen.  
DR Pfam; PF01391; Collagen.1.  
DR ProDom; PD000007; Clg helix.1.  
KW Collagen; Hypothetical protein.  
SQ SEQUENCE 245 AA; 26460 MW; 14951B87D8181A0E CRC64;  
Query Match 98.9%; Score 1292; DB 2; Length 245;  
Best Local Similarity 98.8%; Pred. No. 9.9e-107;  
Matches 242; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MHPOGRAAPQQLLGLFVLLLLQLSAPSSASNPVKQKALIRQREVVDLYNMCILQG 60  
DB 1 MHPOGRAAPQQLLGLFVLLLLQLSAPISASNPVKQKQALLIRQREVVDLYNMCILQG 60  
QY 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGRKGEKGECLRESFESWTPTNTKQCSWSLNYGI 120  
DB 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGRKGEKGECLRESFESWTPTNTKQCSWSLNYGI 120  
QY 121 DLGKTAECTFTGRNSALRVLFSGSLRLKCRNACCQWYFFNGACSGPLPIATITL 180  
DB 121 DLGKTAECTFTGRNSALRVLFSSSLRLKCRNACCQWYFFNGACSGPLPIATITL 180  
QY 181 DQGSPELNSTINIRHTSSVEGLCEGIGAGLVDAVIAWGTGSDPYKGDASTGNVSRIII 240  
DB 181 DQGSPELNSTINIRHTSSVEGLCEGIGAGLVDAVIAWGTGSDPYKGDASTGNVSRIII 240  
QY 241 EELPK 245  
DB 241 EELPK 245  
RESULT 3  
Q96CG8 PRELIMINARY; PRT; 243 AA.  
ID Q96CG8  
AC Q96CG8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DR 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Collagen triple helix repeat containing 1 (collagen triple helix  
DE repeat-containing protein 1).  
GN Name=CTRNC1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
RA Altschul S.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueffing T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Lonellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA Krzywinski M.I., Skalski U., Smallus D.E., Scherch R.A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human

Query	Match	93.7%	Score 1225	DB 2	Length 243
Best Local Similarity	94.7%	Pred. No.9-101			
Matches 232	Conservative	3	Missmatches 8	Indels 2	Gaps 1
QY	1 MHPOGRAASPOLLLFLVLLLLTSLAPSSASENPKYKOKLIROREVVDLYNGMCLOG	60			
DB	1 MRQGPAPASPORLRC-LILLULLLPAPSSASELPKQKQKQLQREVVVDLYNGMCLOG	58			
QY	61 PAVGVRDQSGPANGIPGTPGIPGRDGFKEGKECELRSPESWTPNPKQCSWSSLYNGI	120			
DB	59 PAVGVRDQSGPANGIPGTPGIPGRDGFKEGKECELRSPESWTPNPKQCSWSSLYNGI	118			
QY	121 DLGKIAECTFTMRNSALRVLPFSSSLKAKCNACCQRYFFFNAGESGPIPIAIIYL	180			
DB	119 DLGKIAECTFTMRNSALRVLPFSSSLKAKCNACCQRYFFFNAGESGPIPIAIIYL	178			
QY	181 DOGSFELSTNINHHTSSVEGCEGIGAGLVAVLWVGCSPDPKGDASTGNNSVRII	240			
DB	179 DOGSFELSTNINHHTSSVEGCEGIGAGLVAVLWVGCSPDPKGDASTGNNSVRII	238			
QY	241 EELPK 245				
DB	239 EELPK 243				
RESULT 4					
Q6WM91	PRELIMINARY:	PRT:	243 AA.		
ID	Q6WM91				
AC	Q6WM91				
DT	05-JUL-2004 (Tremblrel. 27, Created)				
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)				
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)				
DE	CTHRC1.				
GN	ORFNames=UNO762;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RA	SEQUENCE FROM N.A.				
RA	MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;				
RA	Clark H.F., Gurney A.L., Adaya E., Baker K., Baldwin D., Brush J.,				
RA	Chen Y., Chow B., Chui C., Crowley C., Currie B., Deuel B., Dowd P.,				
RA	Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heldens S.,				
RA	Huang A., Kim H.S., Klimowaki L., Jin Y., Johnson S., Lee J.,				
RA	Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,				
RA	Seehagiri R., Simmons L., Singh J., Smith V., Stinson J., Vagte A.,				
RA	Vandlen R., Watanabe C., Wrand D., Woods K., Xie M.H., Yansura D.,				
RA	Yi S., Yu G., Yuan J., Zhang W., Zhang Z., Goddard A., Wood W.I.,				

R	A		Godowaki P.; write secreted protein discovery initiative (SDPI), a large-scale effort to identify novel human secreted and transmembrane proteins; a bioinformatics assessment.";
R	T		bioinformatics assessment."/;
R	L		Genome Res. 13:2265-2270(2003).
D	R		EMBL; AY389273.1; "-.
D	R		GO; GO:0005797; C:cytoslaem; IEA.
D	R		GO; GO:0006817; P:phosphate transport; IEA.
D	R		IInterPro; IPR008161; Clg_helix.
D	R		IInterPro; IPR008160; Collagen.
D	R		PProdum; PD000007; Clg_helix; 1.
SQ	KW		COLLAGEN.
SQ	SEQUENCE	243 AA; 26266 MW; BCB49AF4DBC303BC CRC64;	
QY	Query Match	93.0%; Score 1216; DB 2; Length 243;	
Bd	Best Local Similarity	94.3%; Pred. No. 5.7e-100;	
Matches	231; Conservative	3; Mismatches 9; Indels 2; Gaps 1;	
Db			
QY	61 PAGVGRDGSBGANGIPGTPGIRPDGPKEGECELRSEPSWTPTYKCSWSLNYGI	120	
Db	59 PAGVGRRDGSGANAVITPGTPTIGRPDRDPKGEGECLRESFESMTPNVKCSWSLNYGI	118	
QY	121 DLGLIACTFTPKRNSNALRYLPBGGSLRLKCRRNAACCRWRYYTFNGABCGPLPEAIYYL	180	
Db	119 DLGLIACCTFTPKRNSNALRYLFSSLSRLKCRRNACCORWFYTFNGABCGSPLPEAIYYL	178	
QY	181 DQSPPNSTNTIHRTSVSELCAGTAGLVDAIVWGTCSDPYRGDASTGMNVSRITII	240	
Db	179 DQSPPNMSTNIHRTSVSELCAGTAGLVDAIVWGTCSDPYRKGDASTGMNVSRITII	238	
QY	241 EELPCK 245 		
Db	239 EELEPC 243		
RESULT 5			
ID	QBIX63 PRELIMINARY; PRY; 232 AA.		
AC	QBIX63;		
DT	01-MAR-2003 (TREMBLrel_23, Created)		
DT	01-MAR-2003 (TREMBLrel_23, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel_25, Last annotation update)		
DE	NTRYC1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
SP	SEQUENCE FROM N.A.		
RP	Sanuki N., Fujiki K., Kanai A., Tanaka Y., Iwata T.;		
RL	Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; AF395488; AAC17919.1; "		
DR	GO; GO:0005737; C:citosplaem; IEA.		
DR	GO; GO:0006817; P:phosphate transport; IEA.		
DR	IInterPro; IPR008161; Clg_helix.		
DR	IInterPro; IPR008160; Collagen.		
DR	Pfam; PF01391; Collagen; 1.		
DR	PProdum; PD000007; Clg_helix; 1.		
KM	Collagen.		
SO	SEQUENCE 232 AA; 25163 MW; B9DABC30304837ED CRC64;		
QY	Query Match	81.3%; Score 1063; DB 2; Length 232;	
Bd	Best Local Similarity	87.2%; Pred. No. 2.2e-86;	
Matches	197; Conservative	7; Mismatches 12; Indels 10; Gaps 1;	
Db			
QY	29 PSSASNPXYKKAKALIFQR-----VVLDYNMGCIQGPAVGVRGDSFGANGIPG	78	
Db	3 PGGRISIVKREKVKSKLNNGBSAFOGLICGYTNMCIQAPGVGRDSSFANGIPG	62	



RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Eye;  
 RC Strausberg R.;  
 RA Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC040774; AAH40774.1; -  
 DR HSSP; Q60994; IC28.  
 DR MGD; MGI:3032521; BC040774.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR CO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR001073; Clq.  
 DR InterPro; IPR008160; Clqagen.  
 DR InterPro; IPR008983; TNF\_like.  
 DR Pfam; PF00386; Clq; 1.  
 DR Pfam; PF01391; Clqagen; 1.  
 DR PRINTS; PR00007; COMPLEMENTC1Q.  
 DR SMART; SM00110; Clq; 1.  
 DR PROSITE; PS01113; Clq; 1.  
 DR Collagen.  
 KW KW  
 SQ SEQUENCE 287 AA; 29292 MW; 8FF89EC1C7420415 CRC64;  
 Query Match 10.2%; Score 133; DB 2; Length 287;  
 Best Local Similarity 27.5%; Pred. No. 0.0015;  
 Matches 46; Conservative 16; Mismatches 61; Indels 44; Gaps 3;  
 QY 12 LILGLFLVLLILQLSPSSASSENPKYKOKALRQREVVDLYN----- 54  
 Db 1 MALGLLAVFLQLQAAPPGA-----HYEMLGTCRMTCDPYVADPAGPAGAKAPPGP 54  
 QY 55 -----GMCLOGPAGVPGRDSPGANGIPGTGIPGRDGFKEGECLE 98  
 Db 55 STAALEWODLSANPPPPFQGPKGDPGRKGRKPPGPPEPPRGPGRGPGSGR 114  
 QY 99 SFESESTPNKQCSWSGLNYGIDLGKIAECTFTFRKMSNGLRYLFSG 145  
 Db 115 GL-----PGLQLTTSAAAGVGVSGGTGGGCTGEVTSALSAFSG 156  
 RESULT 9  
 OX036 PRELIMINARY; PRT; 565 AA.  
 AC OX036;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Collal protein.  
 GN Name=Collal;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_Taxid=10090;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=mx FVB/N; TISSUE=Mammary tumor;  
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=mx FVB/N; TISSUE=Mammary tumor;  
 RC Strausberg R.;  
 RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC034164; AAH34164.1; -  
 DR MGD; MGI:1277201; Coll13a1.  
 DR GO; GO:0005911; C:intercellular junction; IDA.  
 DR InterPro; IPR008161; Clq\_helix.  
 DR InterPro; IPR008160; Clqagen.  
 DR Pfam; PF01391; Clqagen; 5.  
 DR Prodom; PD000007; Clq\_helix; 1.  
 DR Collagen.  
 KW KW  
 SQ SEQUENCE 565 AA; 56726 MW; DBD3FF99D670195F CRC64;  
 Query Match 10.1%; Score 132.5; DB 2; Length 565;  
 Best Local Similarity 32.7%; Pred. No. 0.0036;  
 Matches 36; Conservative 10; Mismatches 35; Indels 29; Gaps 5;  
 QY 27 SAPSSASSENPKYKOKALRQREVVDLYN-----MCLQGPAGVPGRDSPGANG 75  
 Db 402 SGPKGAKGPP-----GKGEWVD-YNGSINEALQEIRTLALMGPPGLPGGTGPPGPG 452  
 QY 76 IP-----GTGIPGRDGFKEGECLESESESTPNKQCSWSGLNYG 119  
 Db 453 TEGQGEIGLPGPPGHDGKGRG---KPGDAGWSRTPRKGMASRNG 499  
 RESULT 10  
 OX6ZM13 PRELIMINARY; PRT; 551 AA.  
 AC OX6ZM13;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein FLJ23917.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Colon;  
 RC Kawabata A., Hikiyi T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Ota T., Suzuki Y., Obaayaishi M., Nishi T., Shibahara T.,  
 RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;  
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK117256; BAD18742.1; -

```
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR003112; Clfac like.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF02191; OLF; 1.
DR Prodom; PDOM00007; Clg_helix; 1.
DR SMART; SM00284; OLF; 1.
DR Collagen.
KW SEQUENCE.
SQ SEQUENCE 551 AA; 58957 MW; CE14A36120DEC18 CRC64;

Query Match 10.0%; Score 130.5; DB 2; Length 551;
Best Local Similarity 30.7%; Pred. No. 0.0052;
Matches 35; Conservative 18; Mismatches 20; Indels 41; Gaps 5;

QY 3 PGRASAPOLLGLFLVLLLLQLSAPSSASEN-----PKVKOK-----AL 43
DB 76 PRGASAPQ-----DPASARAKRSHSGEPAPHIAESHDMMMWTYSM 119
QY 44 IROREVVDLYN--GMCLGPGAVPGRDGSFGANGIPGTPGIPGRDGFKEKGE 94
DB 120 VPIRVAVDLCNSTKGICLTGSPGSP--GPPGAGGLPGHNGLDGDPGPGKGE 170

RESULT 11
ID Q9QZ50 PRELIMINARY; PRT; 1669 AA.
AC Q9QZ50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha 3 collagen IV.
OS Name=Col14a3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
MEDLINE=20005934; PubMed=10534397; DOI=10.1006/geno.1999.5943;
RA Lu W., Phillips C.L., Killen P.D., Hlaing T., Harrison W.R.,
RA Bider F.F.B., Miner J.H., Overbeek P.A., Meisler M.H.;
RT "Insertional mutation of the collagen genes Col4a3 and Col4a4 in a
RT mouse model of Alport syndrome.";
RL Genomics 61:113-124(1999).
DR EMBL; AF169387; AAD50449.1; -.
DR PIR; I48302; I48302.
DR HSSP; P02462; 1L11.
DR MGD; MGI:104688; Col4a3.
DR GO; GO:0005587; C:collagen type IV; ISS.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005178; F:integrin binding; ISS.
DR GO; GO:0008191; F:metalloendopeptidase inhibitor activity; ISS.
DR GO; GO:0006919; P:caspase activation; ISS.
DR GO; GO:0008283; P:cell proliferation; ISS.
DR GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; ISS.
DR GO; GO:0006917; P:induction of apoptosis; ISS.
DR GO; GO:0016525; P:negative regulation of angiogenesis; ISS.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR004828; Surface_antigen.
DR InterPro; IPR001442; Procollagen_C.
DR Pfam; PF01413; C4; 2.
DR Pfam; PF01391; Clg_helix; 9.
DR Prodom; PDOM00007; Clg_helix; 2.
DR Prodom; PDI53432; Surface_antigen; 2.
DR Prodom; PD003923; Procollagen_C; 2.
DR SMART; SM00111; C4; 2.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW Collagen.
```

```
SQ SEQUENCE 1669 AA; 161769 MW; 30976E59739A47B2 CRC64;

Query Match 9.9%; Score 129.5; DB 2; Length 1669;
Best Local Similarity 37.0%; Pred. No. 0.022;
Matches 37; Conservative 9; Mismatches 29; Indels 25; Gaps 4;

QY 1 MHQGRASPOLLLGLFLVLLLLQLSAPSSASENPKVKOKALIROREVVDLYNMGCT-- 58
DB 1 MHSK---TAPRL--VFLLITLLLLAASPVAASKCVCVKGK-----GQCLCA 42
QY 59 -----QGPAGVPGRDSPGANGIPGTPGIPGRDGFKEKGE 93
DB 43 GTKGEKGEKGVGSPGPGQKGFPGEGLPQPGKSGPG 82

RESULT 12
ID Q8BVD7 PRELIMINARY; PRT; 289 AA.
AC Q8BVD7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 16 days embryo lung cDNA, RIKEN full-length enriched
DE library, clone:8430425G24 product:COMPLEMENT-C1Q TUMOR NECROSIS
DE FACTOR-RELATED PROTEIN homolog.
GN Name=C1qtnf7;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,
RA Komoto H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Komoto H., Akiyama J., Nishi K., Kitsuana T., Teshiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiwa K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa K., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
```



[illegible]

```

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein ZK1010.7.
GN ORFNames=ZK1010.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RL investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z82083; CAB04973.1; -.
DR PIR; T27644; T27644.
DR WormBase; WBGene0000672; ZK1010.7.
DR WormPep; ZK1010.7; CE15507.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0042302; F:structural constituent of cuticle; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002486; Col_cuticle_N.
DR Pfam; PF01484; Col_cuticle_N; 1.
DR Collagen; Hypothetical protein.
SQ SEQUENCE 298 AA; 30260 MW; BE5FD5B5D5C635DD CRC64;

Query Match          9.64; Score 125.5; DB 2; Length 298;
Best Local Similarity 39.54; Pred. No. 0.0073;
Matches 30; Conservative 6; Mismatches 29; Indels 11; Gaps 2;

QY 31 SASENPKVQKALIRQREVVDLYNGMCLQGPGVPGHSGPGANGIPGTGIPGRDQFKG 90
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 82 SSNETTRLKRDAACQ-----GCCL---PGLPGDPGPPGKXGAPGRGAPGAPQFP 130
   |||||

QY 91 EKGECLRESFESWTP 106
   |||||
Db 131 RPPAVCEITEPPCTP 146
   |||||

```

Search completed: August 4, 2005, 04:21:22  
 Job time : 93.3529 secs